MiraiBio

DNASIS® MAX

Version 2.5

User's Manual

LICENSE AGREEMENT

BEFORE OPENING THIS PACKAGE, YOU SHOULD CAREFULLY READ THE FOLLOWING TERMS AND CONDITIONS. BY OPENING THIS PACKAGE YOU AGREE TO BECOME BOUND BY THE TERMS AND CONDITIONS OF THIS AGREEMENT, WHICH INCLUDES THE SOFTWARE LICENSE AND LIMITED WARRANTY. IF YOU DO NOT AGREE WITH THESE TERMS AND CONDITIONS, YOU SHOULD PROMPTLY RETURN THE PACKAGE UNOPENED TO MIRAIBIO, INC. ("Mirai") or Mirai Distributor AND YOUR MONEY WILL BE REFUNDED.

The enclosed software is licensed, not sold, to you for use only upon the terms of this Agreement, and Mirai reserves any rights not expressly granted to you. You are responsible for the selection of the Software to achieve your intended results and for the installation, use and results obtained from the Software. You own the media on which the Software is originally or subsequently recorded or fixed, but Mirai retains ownership of all copies of the Software itself.

LICENSE

You may:

a. Use the Software on a single machine at any given time.

b. Obtain limited numbers of Copy Protection Devices. Additional, Copy Protection Devices are provided only as a convenience of running the software.

c. In no manner engineer or reverse-engineer the copy protection hardware, or whole or part of the software.

d. Copy the software only for backup provided that you reproduce all copyright and other proprietary notices that are on the original copy of the Software provided to you. Certain Software, however, may include mechanisms to limit or inhibit copying. Such Software is marked copy protected.

e. Transfer of the Software and all rights under this Agreement to another party together with a copy of this Agreement if the other party agrees to accept the terms and conditions of this Agreement. If you transfer the Software, you must at the same time either transfer all copies whether in printed or machine-readable form, to the same party or destroy and copies not transferred.

RESTRICTIONS

You may not use, copy, modify, or transfer the Software, or any copy, in whole or in part, except as expressly provided for in this Agreement. Any attempt to transfer any of the rights, duties or obligations hereunder except as expressly provided for in this Agreement is void. YOU MAY NOT RENT, LEASE, LOAN, RESELL FOR PROFIT, OR DISTRIBUTE.

TERM

This Agreement is effective until terminated. You may terminate it at any time by destroying the Software together with all copies in any form. This Agreement will immediately and automatically terminate without notice if you fail to comply with any term or condition of this Agreement. You agree upon termination to promptly destroy the Software together with all copies in any form.

LIMITED WARRANTY

Mirai warrants, for the period of ninety (90) days from the date of delivery of the Software to you as evidenced by a copy of your receipt, that:

(1) The Software, unless modified by you, will perform the function described in the documentation provided by Mirai.   Your sole remedy under the warranty is that Mirai will undertake to correct within a reasonable period of time any marked Software Error (failure of the Software to perform the functions described in the documentation). Mirai does not warrant that the Software will meet your requirements, that operation of the Software will be uninterrupted or error-free, or that all Software Errors will be corrected.

(2) The media on which the Software is furnished will be free from defects in materials and workmanship under normal use.   Mirai will, at its option, replace or refund the purchase price of the media at no charge to you, provided you return the faulty media with proof of purchase to Mirai. Mirai will not have any responsibility to replace or refund the purchase price of the media damaged by accident, abuse or misapplication.

THE ABOVE WARRANTIES ARE EXCLUSIVE AND IN LIEU OF ALL OTHER WARRANTIES, WHETHER EXPRESS OR IMPLIED, INCLUDING THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. NO ORAL OR WRITTEN INFORMATION OR ADVICE GIVEN BY MIRAI, ITS EMPLOYEES, DISTRIBUTORS, OR AGENTS SHALL INCREASE THE SCOPE OF THE ABOVE WARRANTIES OR CREATE ANY NEW WARRANTIES. SOME STATES DO NOT ALLOW THE EXCLUSION OF IMPLIED WARRANTIES, SO THE ABOVE EXCLUSION MAY NOT APPLY TO YOU. IN THAT EVENT, ANY IMPLIED WARRANTIES ARE LIMITED IN DURATION TO NINETY (90) DAYS FROM THE DATE OF DELIVERY OF THE SOFTWARE. THIS WARRANTY GIVES YOU SPECIFIC LEGAL RIGHTS. YOU MAY HAVE OTHER RIGHTS, WHICH VARY FROM STATE TO STATE.

LIMITATIONS OF REMEDIES

Mirai's entire liability to you and your exclusive remedy shall be the replacement of the Software media or the refund of your purchase price as set forth above. If Mirai or the Mirai's distributors are unable to deliver replacement media which is free of defects in materials and workmanship, you may terminate this Agreement by returning the Software and your money will be refunded.

REGARDLESS OF WHETHER ANY REMEDY SET FORTH HEREIN FAILS ITS ESSENTIAL PURPOSE, IN NO EVENT WILL MIRAI BE LIABLE TO YOU FOR ANY DAMAGES, INCLUDING ANY LOST PROFITS, LOST DATA OR OTHER INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY OF SUCH DAMAGES, OR FOR ANY CLAIM BY ANY OTHER PARTY.

SOME STATES DO NOT ALLOW THE LIMITATION OR EXCLUSION OR LIABILITY FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES TO THE ABOVE LIMITATION OR EXCLUSION MAY NOT APPLY TO YOU.

GOVERNMENT LICENSEE/RESTRICTED RIGHTS LEGEND

If you are acquiring the Software on behalf of any unit or agency of the United States Government, the following provisions apply.

The Government acknowledges Mirai's representation that the Software and its documentation were developed at private expense and no part of them is in the public domain.

The Government acknowledges Mirai's representation that the Software is Restricted Computer Software as that term is defined in Clause 52.227-19(a) of the Federal Acquisition Regulations (FAR).

The Government acknowledges that the Software is classified as Commercial Computer Software and the Government is acquiring only restricted rights in the Software and its documentation will be as defined in Clause 52.227-19 (c)(1) and (2) of the FAR.

Manufacturer is MiraiBio, Inc., 1201 Harbor Bay Parkway, Suite 150, Alameda, CA 94502

EXPORT LAW ASSURANCES

You acknowledge and agree that the Software is subject to restrictions and controls imposed by the United States Export Administration Act ("The Act") and the regulations thereunder. You agree and certify that neither the Software nor any direct product thereof is being or will be acquired, shipped, transferred or reexported, directly or indirectly, into any country prohibited by the Act and the regulations thereunder or will be used for any purpose prohibited by the same.

GENERAL

This agreement will be governed by the laws of the State of California, except for that body of law dealing with conflicts of law.

Future updates of the Software will be available for purchase by licensees for a fee provided a registration card has been received by MiraiBio, Inc.

Should you have any questions concerning this Agreement, you may contact Mirai at http://www.miraibio.com.

You acknowledge that you have read this Agreement, understand it and agree to be bound by its terms and conditions. You further agree that it is the complete and exclusive statement of the agreement between us which supercedes any proposal or prior agreement, oral or written, and any other communications between us in relation to the subject matter of this Agreement.

Contents

Preface

DNASIS MAX v2.5 USER'S MANUAL

Published by:

Hitachi Software Engineering Co., Ltd.

Address:

    Hitachi Software Engineering Co., Ltd.

    Life Science Research Center

    1-1-43 Suehiro-cho Tsurumi-ku

    Yokohama, 230-0045 Japan

First Edition November 2001 (invalid)

Second Edition February 2003 (invalid)

Third Edition November 2003 (invalid)

Fourth Edition March 2004

©November 2001 Hitachi Software Engineering Co., Ltd.   All rights reserved.

DNASIS® is a registered trademark of Hitachi Software Engineering Co., Ltd.

Windows® is a registered trademark of Microsoft Corporation. NCBI and BLAST are software products developed by the National Center for Biotechnology Information. Primer3 is a software product developed by the Whitehead Institute for Biomedical Research. All other company and product names mentioned in this manual are trademarks or registered trademarks of their owners.

Under the approval of UK Medical Research Council, our waveform display program uses the io_lib library developed by Staden Package of the U.K. The Multiple Alignment method uses the EMBL-licensed ClustalW. The plasmid mapping function uses the library of Rogue Wave Stingray Studio, and with the consent of Rogue Wave Software, Inc.

Technical Support Information

United States

MiraiBio, Inc.

1201 Harbor Bay Parkway Ste. 150

Alameda, CA 94502

USA Only: 1-800-624-6176

Tel: +1-510-337-2000

Fax: +1-510-337-2099

www.miraibio.com

gene@miraibio.com

Europe

Hitachi Software Engineering Europe AG

Neues Kranzler Eck

Kurfürstendamm 22

10719 Berlin

Germany

Tel: +49-30-8877-2600

FAX: +49-30-8877-2610

www.hitachisoft-bio.com

info-bio@hitachisoft.de

Japan

Hitachi Software Engineering Co., Ltd.

Life Science Research Center

1-1-43 Suehiro-cho Tsurumi-ku

Yokohama, 230-0045 Japan

TEL: +81-45-500-5111

FAX: +81-45-500-5119

www.hitachisoft.jp

dnasis@hitachisoft.jp

Chapter 1 Window Descriptions

1.1 Main Window

This section explains the Main window of DNASIS MAX.
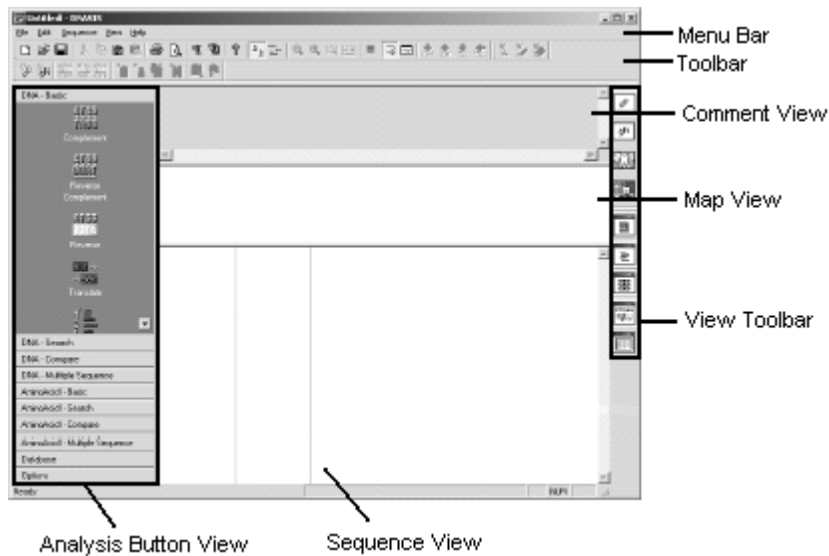
1.2 Description of Individual components of Main Window

Sequence View Pane

The Sequence View Pane displays chromatograms, sequence, the results of sequence editing and the results of analysis.

(1) Indicates the sequence name.

(2) Indicates the analysis name.

(3) Indicates the sequence
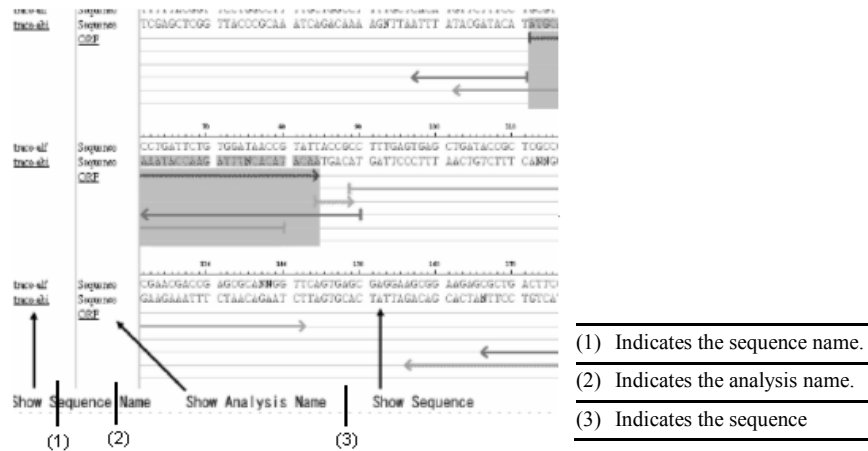
Map View Pane

The Map View Pane provides a map-style overview of the result of analysis currently displayed in the Sequence View.

Using and on the toolbar allows you to expand and shrink images.
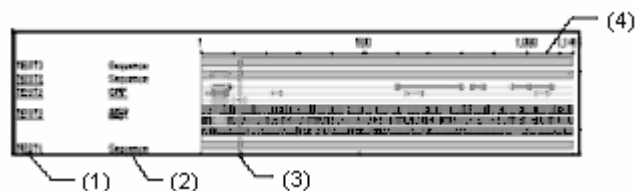
*For details, refer to "About the Target" in "2.9 Editing and Analyzing Multiple Sequences".

(1)	Indicates the sequence name. The sequence specified as the target* in the Sequence View is underlined.
(2)	Indicates the analysis name. The sequence specified as the target* in the Sequence View is underlined.
(3)	Displays the red-framed area in the Sequence View.
(4)	Displays the ruler.
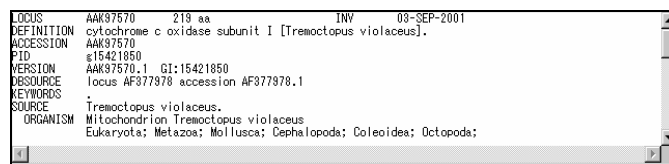
Comment View Pane

The Comment View Pane displays a comment when you read a comment-based file in any of the FASTA, GenBank Flat, EMBL, PIR, and former DNASIS formats.

You can toggle (view and hide) a comment by clicking the button () on the View Toolbar. You can edit the comment directly.

If several sequences have been read and displayed, the comment activated is the one belonging to sequence active in the Sequence View Pane. .

Analysis Button View Pane

The Analysis Button View Pane displays and gives you access to the analysis functions available to you in DNASIS MAX.

Each analysis category is indicated by a tab button and clicking the analysis category name displays the sub menu of functions in that category.
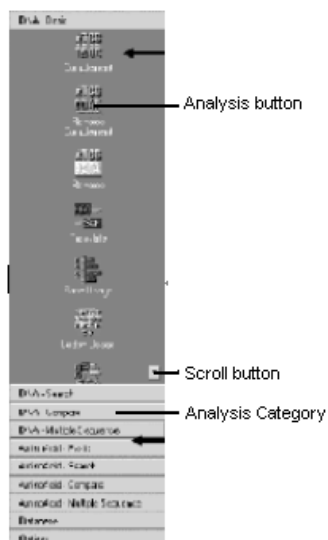
Right-clicking the icon displays a menu as shown below.
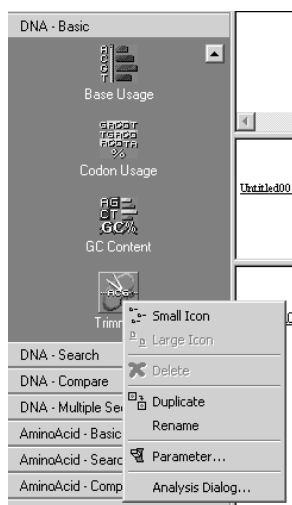
Menu name	Function
Small Icon	Displays a shrunken icon in the leftmost corner of the view.
Large Icon	Displays an enlarged icon in the middle of the view.
Delete	Deletes the analysis icon.
Duplicate	Copies the analysis menu.
Rename	Changes the analysis name.
Parameter...	Displays the Parameterset Editor for changing the setting.
Analysis Dialog...	Starts an Analysis dialog.

1.3 Toolbars

Switch Pane Toolbar

The Switch Pane Toolbar is for controlling and switching between the different views in the Main Window.
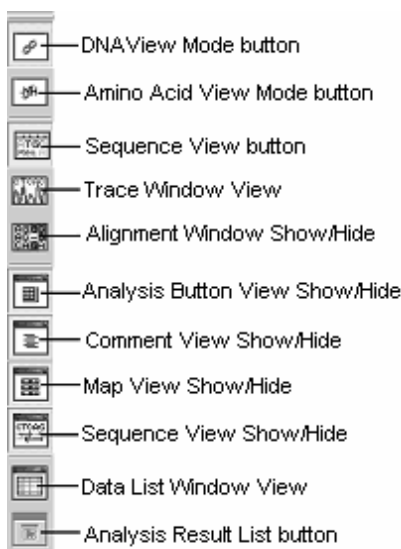
Icon	Function
DNA View Mode button	Shows DNA in sequence view or in alignment view.
Amino Acid View Mode button	Shows amino acids in sequence view or in alignment view.
Sequence View Mode button	Shows sequence in sequence view pane. If you click this button while in trace view or alignment view, the mode will switch to sequence view.
Trace Window View	If this button is on and a trace file available a chromatogram will be seen. The DNA and amino acid toolbars will be hidden.
Alignment Window Show/Hide	If this button is on, the alignment window will be shown.
Analysis Button View Show/Hide	If this button is on, the analysis button view will be shown.
Comment View Show/Hide	If this button is on, the comment view will be shown.
Map View Show/Hide	If this button is on, the map view will be shown.
Sequence View Show/Hide	If this button is on, the sequence view will be shown.
Data List Window View	Click this button to show the data list window.
Analysis Result List button	Click this button to show the analysis result list.

Other Toolbars

Toolbars provide icons for frequently used functions

Standard Toolbar

Icon	Function
	Opens a new empty window.
	Opens a sequence file or project.
	Saves a project disk.
	Cuts a portion of a sequence onto the Clipboard.
	Copies a portion of a sequence onto the Clipboard.
	Copies the image of a displayed view onto the Clipboard.
	Pastes any data from the Clipboard into designated space.

Icon	Function
	Prints the contents of the sequence view pane
	Print Preview
	Display the Preferences window[1].
	Displays the Internet Setting dialog box[2].
	Displays online help.
	Displays the Analysis button using a large icon.
	Displays the Analysis button using a small icon.
	Shrinks the Map View.
	Expands the Map View.
	Expands any selected area in the Map View.
	Displays the Map View at 100% size.
	Displays a sequence in the Sequence View or the result of analysis on a single line.
	Displays a sequence in the Sequence View or the result of analysis by allowing wrapping according to the width of the window. A change in the window size will automatically change the fold-back "wrapping" width accordingly.
	Displays a sequence in the Sequence View or the result of analysis by folding it back according to the number of characters.
	Lowers the order of a sequence.
	Raises the order of a sequence.
	Lowers the order of an analysis.
	Raises the order of an analysis.

DNA Toolbar

Icon	Function
	Adds DNA sequence.
	Shows a translated amino acid sequence that was selected in DNA sequence view in amino acid sequence view.
	Converts into a reverse complement sequence.
	Converts into a complement sequence.
	Converts into a reverse sequence.
	Converts the expression of a sequence into uppercase characters.
	Converts the expression of a sequence into lowercase characters.
	Converts the expression of a sequence with uppercase and lowercase characters.
	Masks the specified or highlighted stretch of sequence..
	Searches a sequence.
	Moves the cursor to the specified position.

Amino Acid Toolbar

Icon	Function
	Adds an amino acid sequence.
	Converts highlighted or entire sequence into uppercase characters.
	Converts highlighted or entire sequence into lowercase characters.
	Converts highlighted or entire sequence to opposite case characters.
	Masks the sequence of a range.
	Searches for a specified sequence string
	Moves the cursor to specified position.

Annotation Toolbar

Icon	Function
	Adds a new annotation.
	Adds an annotation entry.
	Adds more than one annotation entry at the same time.
	Adds a part to an annotation.
	Moves the selected annotation (or part) one step up.
	Moves the selected annotation (or part) one step down.
	Moves the selected annotation (or part) to the top layer.
	Moves the selected annotation (or part) to the bottom layer.

Waveform Toolbar

Icon	Function
	Decreases the vertical width of tracing.
	Increases the vertical width of tracing.
	Decreases the vertical width of a view.
	Decreases the horizontal width of a view.
	Increases the vertical width of a view.
	Increases the horizontal width of a view.
	Turns ON/OFF the hand tool (for scrolling through individual items of data) in the parallel data mode.
A	Views/hides a trace of lane A.
C	Views/hides a trace of lane C.
G	Views/hides a trace of lane G.
T	Views/hides a trace of lane T.
	Selects a reference sequence for alignment (only the FASTA format).
ACGT	Hides an imported sequence.
ACGT ACGT	Switches back to the alignment display mode.
A--T ACGT	Makes an alignment between a trace-indicated sequence and an imported sequence.
ATCG TAGC	Converts into a complement sequence.

1.4 Menu Bar

File menu	Function description
New	Opens a prompt dialog.
Open	Opens a specified project file. It is also possible to specify more than one file at the same time.
Save Project	Stores a project by overwriting it.
Save Project As	Stores a project by giving it a new name.
Export...	Stores a sequence by giving it a name. The file formats below are available. FASTA format Text format Formatted text format MSF format DMP (DNASIS Plasmid Map File) format
Import Sequence...	Obtains a sequence from a file. The target is limited to those items of data that have undergone sequence conversion because of the need for checking the source file for integrity. The target does not cover any file that is incapable of sequence conversion.
Print Setup...	Displays the Set Printing Information window and gives the setting of the paper size and printer information.
Print Preview	Displays a print image.
Print...	Carries out printing.
Print Page Preview	Displays a print image for the part that is currently displayed on the window.
Print Page...	Prints only the part that is currently displayed on the window.
Exit	Terminates the DNASIS MAX.

Edit menu	Function description
Undo	Cancels the previous operation.
Cut	Cuts a sequence portion.
Copy	Copies a sequence portion into the Clipboard.
Copy Image	Copies the image of a displayed view into the Clipboard.
Paste	Pastes any items of data on the Clipboard into a specified part.
Select All	Highlights all the sequence data or comments where the cursor is located.
Select Range...	Highlights a range.

Sequence menu	Function description
NewDNA	Adds a DNA sequence.
New Amino Acid	Adds an amino acid sequence.
Duplicate	Creates a new sequence by duplicating a currently selected sequence.
Revert	Returns to the pre-edit sequence.
Find...	Searches for a sequence.
Find Again	Searches for the next occurrence. It will become possible to select this only after Find is executed.
Jump	Moves the cursor to a specified position.
Complement	Converts into a complement sequence.
Reverse	Converts into a reverse sequence.
Reverse Complement	Converts into a reverse complement sequence.

Upper Case	Converts the expression of a sequence into uppercase characters.
Lower Case	Converts the expression of a sequence into lowercase characters.
Exchange Case	Converts the expression of a sequence with uppercase and lowercase characters.
Mask	Masks the sequence of a range.
Make Consensus	Uses the Editor to enter a consensus sequence as a new sequence in the alignment mode.

View menu	Function description
Analysis Button View	Views/hides the Analysis Button View.
Comment View	Views/hides the Comment View.
Map View	Views/hides the Map View
Data List...	Displays the Data List window.
Standard Toolbar	Views/hides the Standard toolbar.
Switch Pane Toolbar	Views/hides the Window Switchover toolbar.
DNA Toolbar	Views/hides the DNA toolbar.
Amino Acid Toolbar	Views/hides the Amino Acid toolbar.
Annotation Toolbar	Views/hides the Annotation toolbar.
Status Bar	Views/hides the Status Bar.
Preferences...	Displays the Preferences[1].
Internet Options...	Displays the Internet Setting dialog box[2].

[1]For details, refer to "1.5 Preferences Dialog Box".

[2]For details, refer to "1.6 Internet Setting Dialog Box".

Help menu	Function description
Contents	Displays online help.
User Forum Web Page	Displays a Web site for the User Forum of DNASIS MAX. This requires an environment capable of being connected to the Internet.
About DNASIS MAX...	Displays the version information.

Popup Menu on Annotation Display

Menu	Description
Selected Annotation Setting...	Displays the Annotation Setting dialog to set the parameters for the selected annotation (or part).
Annotation List...	Displays the Annotation List dialog to list up annotations.
Kind Color Setting...	Displays the Kind Color Setting dialog to set colors.
New Annotation	Create a new annotation.
Add Annotation To Selected Area	Adds the selected range as an annotation.
Add Annotations To Selected And Emphasized Area	Adds a selected area and an emphasized area as separate annotations.
Add Annotation Parts To Selected And Emphasized Area	Adds annotations that have selected areas and emphasized areas as separate annotation parts.
Duplicate Annotation	Duplicates a selected annotation.
Show All Annotations	Shows all the annotations, including the hidden annotations.
Hide Selected Annotations	Hides selected and emphasized annotations, including partly selected or emphasized ones.
Delete Selected Annotations	Deletes the selected or highlighted annotation (or part).
Move Up Selected Annotation	Moves the selected annotation (or part) one step forward.
Move Down Selected Annotation	Moves the selected annotation (or part) step down.
Move Selected Annotation To Top Layer	Moves the selected annotation (or part) to the top layer.

Move Selected Annotation To Bottom Layer	Moves the selected annotation (or part) to the bottom layer.
Show Annotation Name and Kind	Shows the name and type of annotation.
Hide Annotation Name and Kind	Hides the name and type of annotation.
Show Link	Opens the URL for the annotation in a browser.
Rearrange Annotations	Restores the annotation modified with Move Up Selected Annotation, Move Down Selected Annotation, Move Selected Annotation To Top Layer and/or Move Selected Annotation To Bottom Layer to the condition at the time of import.

1.5 Preferences Dialog Box
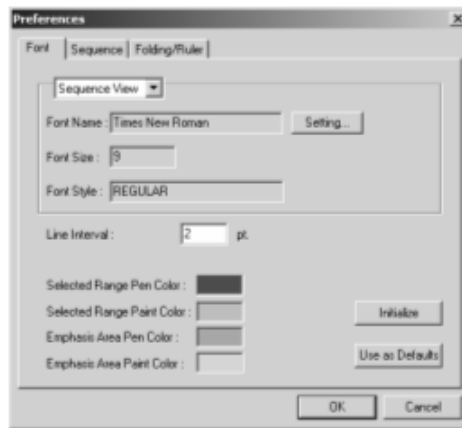
Font Tab

Item	Description
Select View box	Selects a view for use in font setting.
Setting...button	Displays the window for font setting.
Font Name	Indicates the font name.
Font Size	Indicates the font size.
Font Style	Indicates the font style.
Line Interval	Sets the line interval.
Selected Range Pen Color	Sets the color of the selected frame.
Selected Range Paint Color	Sets the color of the selected range.
Emphasis Area Pen Color	Sets the color of the highlighted frame.
Emphasis Area Paint Color	Sets the color of the highlighted range.
Initialize	Initializes all settings to factory presets.
Use as Defaults	Store the current settings as the default settings.

Sequence Tab
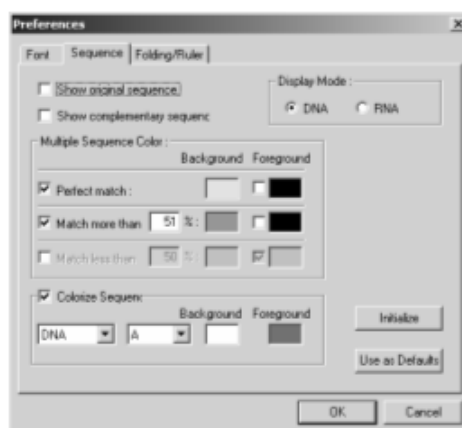
Item	Description
Show original sequence	Displays the pre-edit sequence (the sequence directly after reading from a file) at the same time.
Show complementary sequence	Displays the complement sequence of a sequence at the same time.

Display form of a sequence	Specifies the type (DNA or RNA) of DNA sequence to be displayed.
Emphasis match part in multiple sequence	Sets the background color and character color when displaying alignments.
Perfect match	Sets the color of the background for a position with a perfect match.
Match more than	Sets the color of the background for a position whose hit rate is greater than a specified value.
Match less than	Sets the color of the background for a position whose hit rate is less than a specified value.
Colorize Sequence	When checked, alignments are displayed in color. When not checked, alignments are displayed in black.
Background	Sets the background color for each character. Specify different colors for DNA and amino acids.
foreground	Sets the color for each character. Specify different colors for DNA and amino acids.
Initialize	Initializes all settings to factory presets.
Use Defaults	Stores the settings.

Folding/Ruler Tab
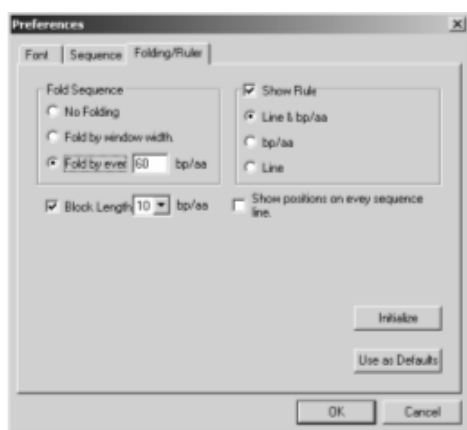
Item	Description
Fold Sequence	Sets how to display sequences in the Sequence View.
No Folding	Uses one line for displaying a sequence.
Fold by window width	Displays a sequence by folding it back according to the width of the window. A change in the window size will automatically change the fold-back width accordingly.
Fold by every bp/aa	Displays a sequence by folding it back according to a specific number of characters.
Block Length:bp/aa	Displays a sequence by inserting a space after a number of characters.
Show Scale	Checking this parameter causes the ruler to be displayed.
Line&bp/aa	Displays both the scale line and the bp indication above a sequence. In the case of alignments, displays the bp count for the consensus sequence.
bp/aa	Displays only the bp indication above a sequence. In the case of alignments, displays the bp count for the consensus sequence.
Line	Displays only the scale line above a sequence.
Show position at the sequence head.	Assigns the bp indication to both the right and left ends of each line of a sequence. For alignments, the value is smaller by the gap.
Initialize	Initializes all settings to factory presets.
Use Defaults	Stores the settings.

1.6 Internet Setting Dialog Box

HTTP Proxy Tab

Item	Description
Server	Specifies the address of a proxy server to connect to the Internet.
Port	Specifies the port number of a proxy server to connect to the Internet.
User Name	Specifies the user name if the proxy server requires user authentication.
Password	Specifies the password if the proxy server requires user authentication.
No Proxy	Specifies a Web address that does not require any connection with a proxy server.
Use Proxy Server	Uses a specified setting to connect to the Internet by way of a proxy server.

FTP Firewall Tab

Item	Description
Server	Specifies the address of a firewall.
Port	Specifies the port number of a firewall.
User Name	Specifies the user name for connection to a firewall.
Password	Specifies the password for connection to a firewall.

Item	Description
Type	Specifies the type of a firewall that is to be used.
Passive Mode	Makes a transfer in the PASV mode.

Mail Tab

Item	Description
Mail Address	Specifies the address of email.
SMTP	Sets the SMTP server for sending messages.
Server	Specifies the name of the SMTP server.
Port	Specifies the port number of the SMTP server.
POP3	Sets the POP server for receiving messages.
Server	Specifies the name of the POP3 server.
Port	Specifies the port number of the POP3 server.
Username	Specifies the account name of the receiving mail server.
Password	Specifies the password of the receiving mail server.

1.7 Data List Window
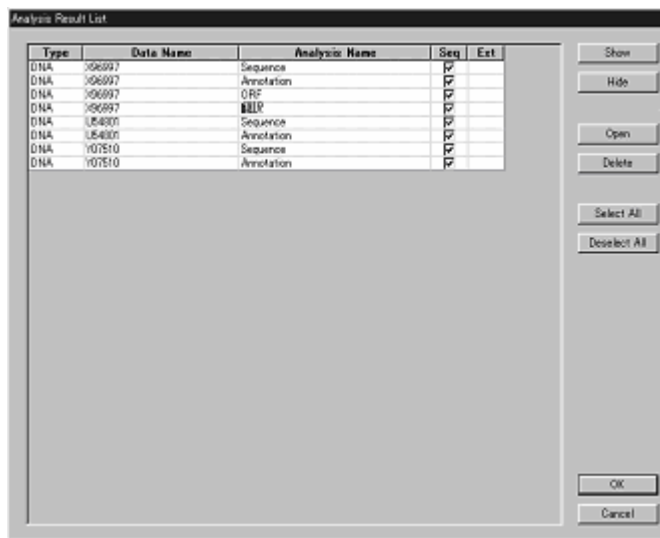
Item	Description
Type	Displays the type of a sequence.
Data Name	Displays the sequence name.
Analysis Name	Displays the analysis name.
Seq	Sets the display condition of the Sequence View.
Ext	Indicates that there is another window indicating the result.
Show	Checks the Seq field for selected analysis.
Hide	Unchecks the Seq field for selected analysis.
Open	Opens the result shown in another window.
Delete	Deletes specified analysis.
Select All	Selects all the lists currently being displayed.
Deselect All	Cancels all the lists currently being displayed.

1.8 Analysis Dialog
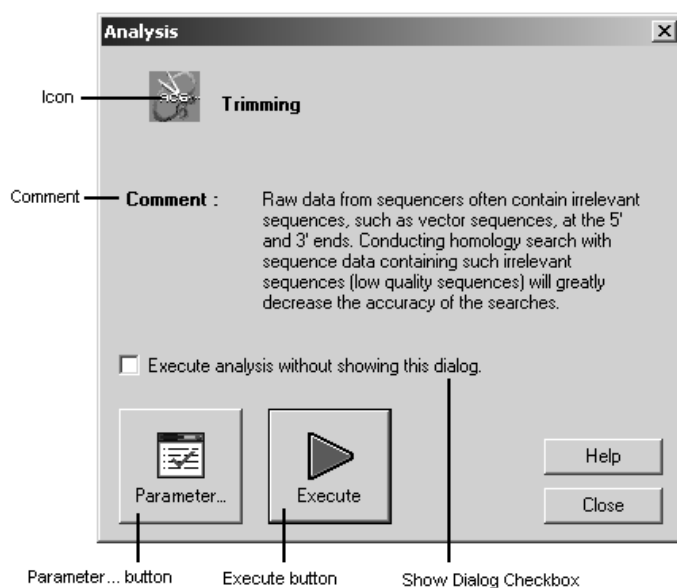
Item	Description
Icon	Shows the analysis button that was selected.
Comment	Shows the content of the analysis button that was selected.
Show Dialog Checkbox	If checked on, from the next time on when the analysis button is clicked analysis will be performed without showing this dialog.
Parameter... button	Starts a parameter dialog.
Execute button	Performs analysis. After analysis the dialog closes.
Help button	Opens the online help.
Close button	Click the button to close the dialog. Settings made in the Analysis dialog are saved.

Chapter 2 DNASIS Basics

2.1 Starting DNASIS

From the Start menu, select the following items: Program, DNASIS MAX, and then DNASIS MAX. The main window will open but behind a dialog box entitled "Welcome to DNASIS MAX".
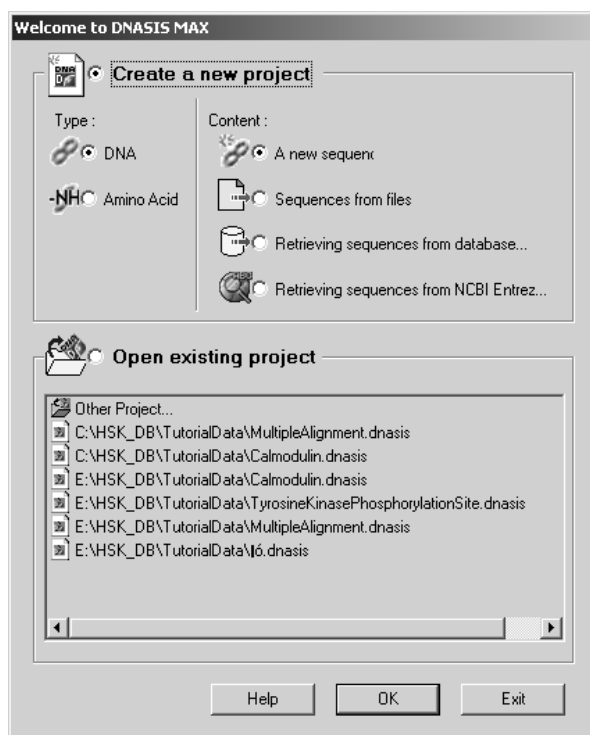
Item	Description
Create a new project button	Creates a new project.
Type button	Sets the sequence type. An error will occur if you specify a different sequence from the type that was set here.
DNA	Sets the sequence type to DNA.
Amino Acid	Sets the sequence type to Protein.
Content button	Sets the sequencing method.
New Sequence	Makes a new DNA or amino acid sequence.
Sequences from files	Imports sequences from a file.
Retrieving sequences from database	Imports sequences from a sequence database. If you select this and click the OK button, a Select Sequence Database dialog will appear.
Retrieving sequences from NCBI Entrez	Obtains sequences from NCBI Entrez. If you select this and click the OK button, an Entrez Search dialog will appear.
Open Existing Project button	Select this radio button if you will open an existing project.
Open project...	Up to 15 recently used projects will appear. To import a project not in the list select Other file… then select the file you want from the file dialog. To import a project not in the list select Other file… then select the file you want from a standard file dialog.
Help button	Opens the online help in a new window.
OK button	Closes the dialog after the parameters have been set with the values entered in the dialog.
Exit button	Closes the dialog without updating the parameters. Closes the application when it is running.

Importing Sequences from a Sequence Database

To start select the Retrieve sequences from database button then click the OK button.

Refer to the sequence database folder specified in the parameters and obtain a database list from that folder.
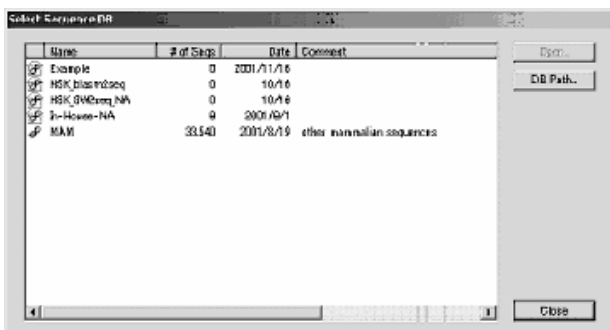
Item	Description
Database List	Shows a sequence database list. The databases listed are only for sequence types (DNA or amino acid) that were specified in the prompt dialog.
Database Icon	Shows the database icon.
Database Name	Shows the database name.
Entries	Shows the number of entries.
Update	Shows the date of update.
Database Comment	Shows comments regarding the database.
Open... button	Click the button to start a Select Entry dialog. Shows the database entry selected from the database list.
DB Path... button	Click the button to start a database path dialog. Specify the sequence database folder to reference.
Close button	Click the button to close the dialog.

Showing Entries

To start click the Entry View button from the Select Sequence Database dialog.
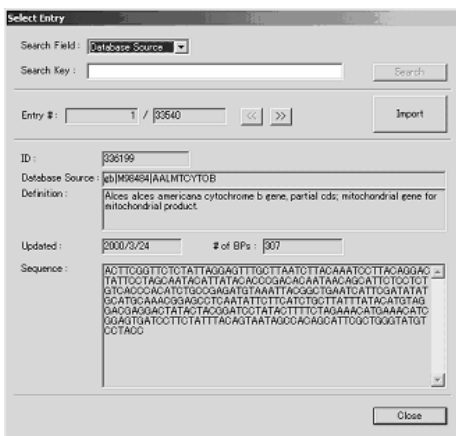
Item	Description
Search Field	Specify the field for the entry search.
Search Key	Input the search key for the entry search. It is possible to enter alphanumerics and symbols.
Search button	Click this button to perform a search. If the search is successful a dialog will show the entry information that was found. If not successful, a message will appear.
Entry	Shows the index number (left) and total number of entries (right) for the entry currently shown.
<< button	If you click the button one entry before the current entry will appear. However, it is not possible to click if the entry currently shown is the first one.
>> button	Click the button to show the entry after the current one. However, it is not possible to click if the entry

	currently shown is the last one.
Import button	Click the button to import the sequence of the entry currently shown into DNASIS.
ID	Shows the ID of the current entry.
Database Source	Shows the Database Source of the current entry.
Definition	Shows the Definition of the current entry.
Updated	Shows the update date of the current entry.
# of BPs	Shows the number of base pairs for the current entry.
Sequence	Shows the sequence of the current entry.
Close button	Click the button to close the dialog.

Obtain sequences from NCBI Entrez.

To start select the Retrieve sequences from NCBI Entrez button then select the OK button.

Item	Description
Search Key Display Textbox	If the search is successful, the search key will appear.
Comment Display Box	If the search is successful, the number of hits will appear.
	If the number of hits and the display number do not match (number of hits exceeds the number specified in the parameter or no invalid search results appear), both numbers will appear.
	If the search is not successful, a message No Result will appear.
Search Result List	If the search is successful, a list of search results will appear. The three display items are Hit Number (with icon), Accession, and Definition. It is possible to sort the list with any of the three items as a key. It is also possible to select multiple items.
	If the search is not successful, nothing will appear in the list.
Search button	Click the button to start the Entrez Search Parameter dialog* and perform a search. When the search is finished the Search Key display textbox, Comment box and Search Result list will be updated.
Import button	Imports the item (GenBank report) selected in the Search Result list into DNASIS.
Close button	Click the button to close the dialog.

For details refer to "3.43 NCBI Entrez Search".

2.2 Entering Sequences

Creating DNA Sequences

1. Select File -> New Menu and a dialog box entitled "Welcome to DNASIS MAX" will appear. For Type select "DNA" and for Content select "A new sequence" then click the OK button.

2. A new untitled Main Window opens with a flashing cursor at position 1in the Sequence view Pane
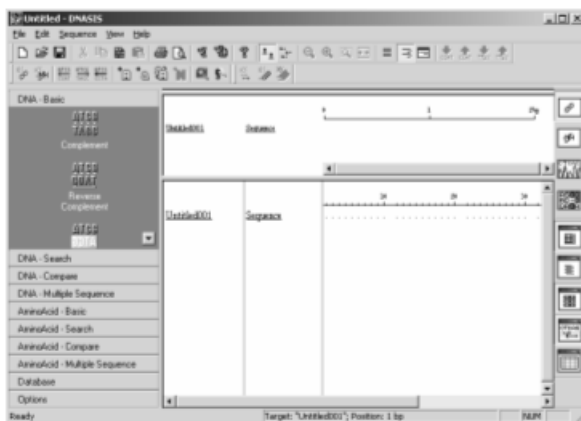
3. Any character entered from the keyboard is inserted at the "Insertion Pointer," which is a vertical bar flashing at the 1bp point in the Sequence View.   You can also paste a sequence from the Clipboard.

Characters You Can Use for DNA Sequences

The following is a list of characters you can enter in DNA sequences.

The input process is case-sensitive; uppercase and lowercase characters are distinguished.   However, the analysis process is not case-sensitive; users are allowed to assign their own meanings to uppercase and lowercase characters.

A
C
G
T
U
R   (Complex code representing A or G)
Y   (Complex code representing C or T)
W   (Complex code representing A or T)
S   (Complex code representing G or C)
K   (Complex code representing G or T)
M   (Complex code representing A or C)
B   (Complex code representing C, G, or T)
D (Complex code representing A, G, or T)
H   (Complex code representing A, C, or T)
V   (Complex code representing A, C, or G)
N
-   (Gap; This can be entered only in the alignment display mode.)

Entering Amino Acid Sequences

1. Select File -> New Menu and a prompt dialog will appear. For Type select Amino Acid and for Content select A new sequence then click the OK button.

2. A new amino acid sequence is produced in the Sequence View.
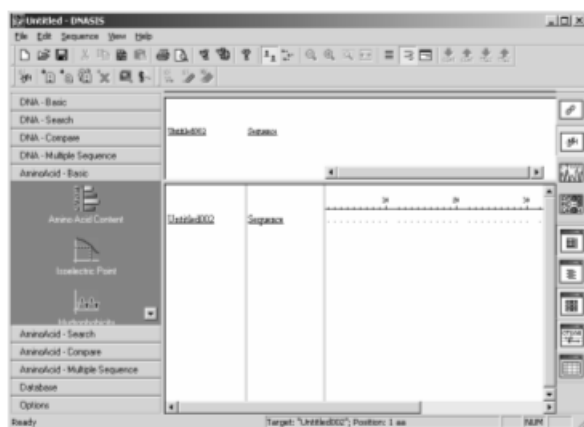
3. Any character entered from the keyboard is inserted at the "Insertion Pointer," which is a vertical bar flashing at the 1aa point in the Sequence View.   You can also paste it from the Clipboard.

Characters You Can Use for Amino Acid Sequences

The following is a list of characters you can enter in amino acid sequences.

Input	Abbreviation	Name
A	Ala	Alanine
R	Arg	Arginine
N	Asn	Asparagine
D	Asp	Aspartic acid
C	Cys	Cysteine
Q	Gln	Glutamine
E	Glu	Glutamic acid
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
L	Leu	Leucine
K	Lys	Lysine
M	Met	Methionine
F	Phe	Phenylalanine
P	Pro	Proline
S	Ser	Serine
T	Thr	Threonine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
V	Val	Valine
B	Asx	Asparagine and aspartic acid
Z	Glx	Glutamine and glutamic acid
*	***	Stop codon (This is displayed at the time of translation from DNA; however, it cannot be entered.)
X	Xxx	Indeterminate amino acid
-		Gap character (only available in alignment view)

The input process is case-sensitive; uppercase and lowercase characters are distinguished. However, the analysis

process is not case-sensitive; users are allowed to assign their own meanings to uppercase and lowercase characters.

Entering and Editing Multiple Sequences

You can edit multiple sequences in a single project window.   With a DNA sequence already displayed, select Sequence and then New DNA; alternatively, you can click the ( ) button on the toolbar.   The new sequence is then added below the existing sequences.

To switch the sequence being edited, click the target sequence.   The same procedure applies to amino acid sequences.
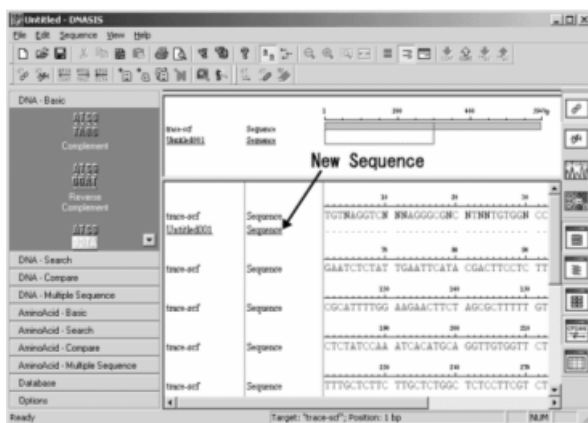
Switching between DNA Sequences and Amino Acid Sequences for Display

DNASIS lets you enter both DNA sequences and amino acid sequences into a single project, although they cannot be displayed in the Sequence view pane at the same time.   It is necessary to switch between DNA sequences and amino acid sequences for display.

To switch to the mode of importing or entering or displaying DNA sequences, click the ( ) button on the View Toolbar.

To switch to the mode of importing, entering or displaying amino acid sequences, click the ( ) button on the View Toolbar.

2.3 Using Existing Files

Opening Sequences from the Menu

You can read sequences from an existing file.

1. Select File -> New Menu and a prompt dialog will appear. For "Type" select DNA and for "Content" select "Sequence from files..." then click the OK button.

2. This displays the file selection dialog box, in which you can select a file or files you want to read.

3. Click the OK button to read the selected files, so that the corresponding sequences are added in the window.

4. It is also possible to select multiple files and read them at once*.
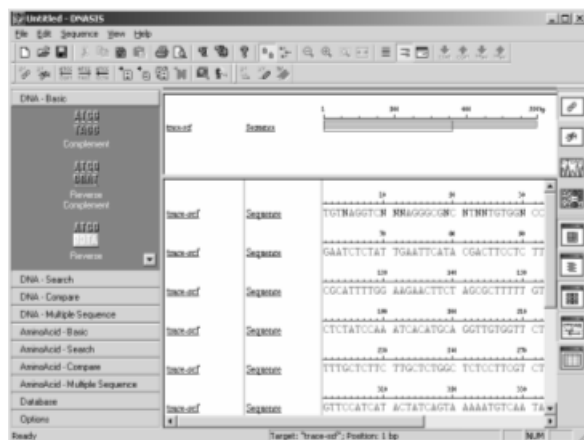
*Refer to "Reading Multiple Files" in "2.3 Using Existing Files".

Opening with the Drag and Drop Method

Using the drag and drop method, you can read files using Windows Explorer.

1. Set the DNASIS Main Window to DNA or Amino Acid mode depending on sequence file type.

2. From Windows Explorer, select a file you want to read.

3. Drag and drop the file in the DNASIS window.
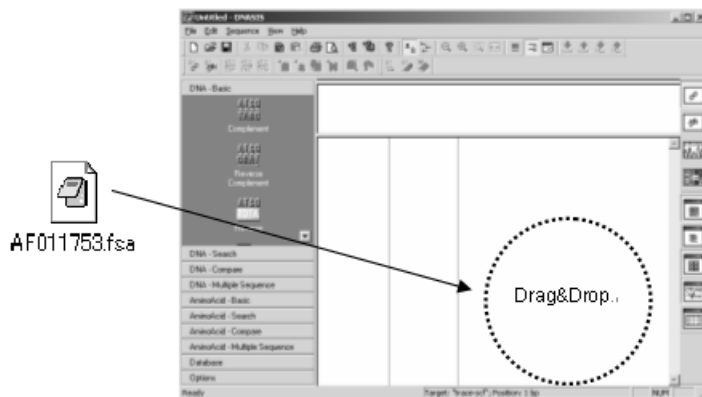
*Refer to "Reading Multiple Files" in "2.3 Using Existing Files".

4. Once dropped, the file is read and the corresponding sequences are added in the window.

5. It is also possible to select multiple files and read them at once*.

Readable File Formats

The following is a list of file formats that can be read and displayed in DNASIS.   These formats are automatically identified by DNASIS according to the contents of files..   You do not need to be concerned about the extensions of file names because they will be ignored.

Format	DNA sequences	Amino acid sequences	Annotation[1]	Features[2]	Trace data[3]
FASTA	ME	ME	ME	NA	NA
GenBank Flat	ME	ME	ME	ME	NA
EMBL	ME	ME	ME	ME	NA
PIR	ME	ME	ME	ME	NA
Former DNASIS	RD	RD	ME	NA	NA
Text	RD	RD	NA	NA	NA
ABI	RD	NA	NR	NA	RD
SCF	RD	NA	NR	NA	RD

RD : Readable

ME : It is possible to read multiple entries from a single file.

NR : Not readable

NA : No applicable data

[1]Annotations can be read in the Comment View. For details, refer to "Comment View" in "1.2 Description of Individual Parts".

[2]By analyzing the "Features" or function descriptions as part of sequences, it is possible to display and edit them in the form of annotations.   For details, refer to "2.11 Annotations".   For EMBL format files, it is possible to show and edit annotations in only "EMBL Nucleotide Sequence Database".   For details about annotations, refer to "2.11 Annotations".

[3]You can display and edit trace data after reading it.   For details, refer to "2.14 Waveform Display Mode".

Reading Files in the FASTA Format

In this format, the entry begins with ">".   The part ranging from line 2 to the point immediately before a line beginning with "//" is read as a sequence   Line 1, ranging from the point after ">" to the end of the line, serves as a sequence name and a comment.   Any sequence including a character that is not found in DNA sequences is regarded as an amino acid sequence; otherwise, the sequence is regarded as a DNA sequence.   If sequences are separated by a "//" line in a file, all of those sequences are read as one.

Reading Files in the GenBank Flat Format

This is a standard format for the GenBank Database.   The part of the Sequence record ranging from the "LOCUS" line to the point immediately before the "ORIGIN" line is displayed as a "comment" in the Comment view pane. The part ranging from the point immediately after the "ORIGIN" line to "//" is displayed as a sequence in the sequence view pane. The first accession number in the ACCESSION line is used as a sequence name.   If there are no accession numbers, the LOCUS name is used as a sequence name.   If the "LOCUS" line includes "aa" as the sequence type, it is regarded as amino acid.   FEATURES are displayed graphically as annotations*.   If sequences are separated by a "//" line in a file, all of those sequences are read into and displayed in one project. In this case a dialog box titled "Import List" will display the list of individual sequences allowing you to select those to be imported.

*Refer to "2.11 Annotations".

Reading Files in the EMBL Format

This is a standard format for the EMBL Nucleotide Sequence Database.   The part ranging from an "ID" line to the point immediately before an "SQ" line is read as a comment, while the part ranging from the "SQ" line to "//" is read as a

sequence.   The first accession number (on the "AC" line) is used as a sequence name.   If there are no accession numbers, the first word on the "ID" line is used as a sequence name.   How to distinguish DNA and amino acid: If the "ID" line includes characters "DNA" or "RNA", it is regarded as DNA.   Also, it is possible to analyze Features and show as annotations.   If sequences are separated by a "//" line in a file, all of those sequences are read into and displayed in one project. In this case a dialog box titled "Import List" will display the list of individual sequences allowing you to select those to be imported.

Reading Files in the PIR Format

This is a standard format for the PIR-International Protein Sequence Database (PIR-PSD).   The part ranging from an "ENTRY" line to the point immediately before a "SEQUENCE" line is read as a comment, while the part ranging from the "SEQUENCE" line to "///" is read as a sequence.   The first word on the "ENTRY" line is used as a sequence name. Any sequence including "#Type Protein" on the "ENTRY line is regarded as an amino acid sequence; otherwise, the sequence is regarded as a DNA sequence.   Also, it is possible to analyze Features and show as annotations.   If sequences are separated by a "//" line in a file, all of those sequences are read into and displayed in one project.

Reading Files in the Old Version DNASIS Format

This format is used for former versions of DNASIS (DNASIS for Windows V2.1 or earlier).   The part ranging from a "DNASIS" line to the point immediately before a "SEQ" line is read as a comment, while the part ranging from the "SEQ" line to "//" is read as a sequence.   The file name without its extension is used as a sequence name.   Any sequence including a character that is not found in DNA sequences is regarded as an amino acid sequence; otherwise, the sequence is regarded as a DNA sequence.

Reading Files in Text Format

Used for text files, this format is different from any of the FASTA, GenBank Flat, EMBL, PIR, and former DNASIS formats.   Excluding the numeric data, symbols, and other characters not found in DNA or amino acid, the entire file is read and displayed as sequence.   If even a single character typically representing an amino acid is present then the sequence is regarded as an amino acid sequence; otherwise, the sequence is regarded as a DNA sequence.   The file name without its extension is used as a sequence name.   The comment is empty.

Reading Trace Data Files in the ABI, SCF and ALF Formats

This deals with trace data files (also referred to as waveforms, chromatograms) in the ABI, SCF and ALF formats. These files typically have the extension .abi, .scf or .alf.   A sequence that has been base-called in advance into the file is extracted as a DNA sequence.   The file name without its extension is used as a sequence name.   The comment is empty.   The trace or waveform data is also extracted and stored in the project and can be viewed along with the sequence in using the "Trace Mode"*.

*Refer to "2.14 Waveform Display Mode".

Reading Multiple Files

You can enter multiple sequences into a single project.   With the DNASIS Main window open, and DNA or Amino Acid mode selected, select File from drop-down menu of Main window and then Open and choose multiple files from the "open files" dialog box and click OK.   Alternatively select File from drop down menu of Main window and then

"Import Sequence". You can also drag and drop a file with multiple sequence records from Windows Explorer.   These

sequences are then added in the window.
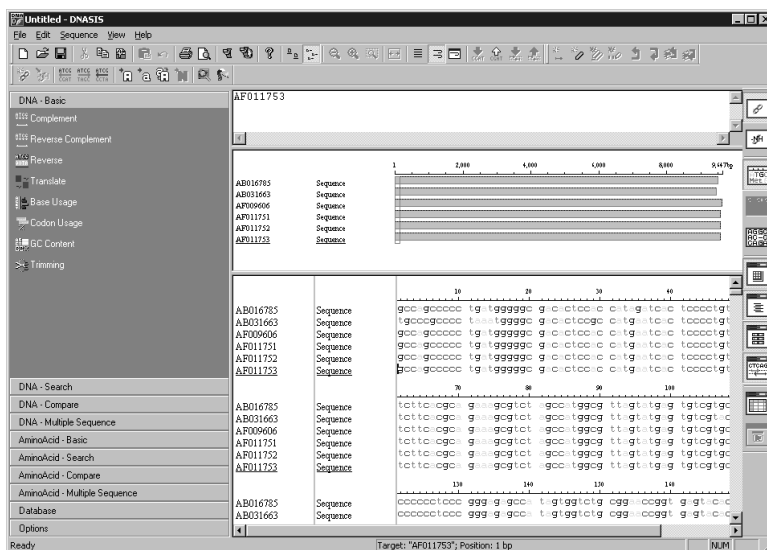
If a file used for import has multiple entries, an entry dialog box will appear where you can import only the entries that

you want.
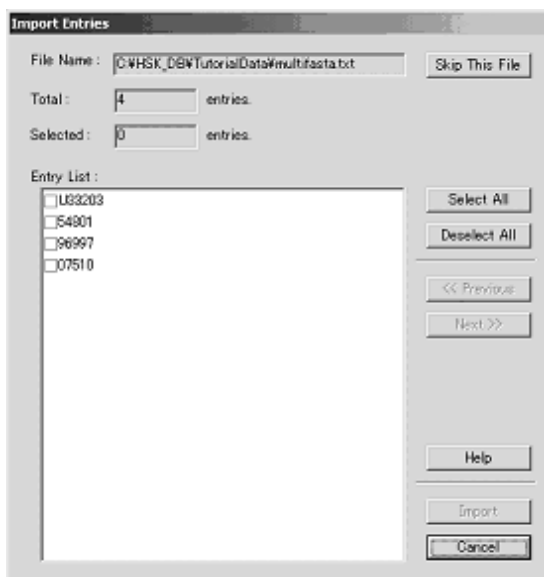
Item	Description
File Name	Shows the name of the currently imported file.
Total	Shows the total number of entries contained in the currently imported file.
Selected	Shows the number of currently selected entries.
Entry List	Shows the entries extracted from a multi-sequence file. An entry with the checkbox on the left selected is an import target. Up to 20 entries appear in one window.
Previous button	Click the button to show the 20 entries before the entry currently shown in the list.
Next button	Click the button to show the 20 entries after the entry currently shown in the list.
Select All button	Selects all entries.
Deselect All button	Unselects all entries.

Help button	Opens the online help.
Import button	Click the button to import the selected entry into the sequence editor and close the dialog.
Skip this file button	Closes the dialog without processing the current file.
Cancel button	Closes the dialog without processing the current file. All subsequent files are also not imported.

About the Sequence Name

Any sequence that has been read from a file is automatically given a sequence name.   For the naming rules, refer to the description of the reading of different file formats into DNASIS.   Be aware that sequence names involve the following restrictions:

Characters that cannot be used:

Japanese Kanji and Kana characters, and the following characters: < > ? * \ : / |

Length: Up to 128 characters.

*Refer to "Renaming Sequences" in "2.4 Editing Sequences (basic)".

Invalid characters for the sequence name are changed into "@" .   A string exceeding the limit of 128 characters will be truncated to the limit.   It is also possible to change sequence names manually*.

The following dialog box appears when there is an attempt to read another sequence of the same name.

Overwrite	Overwrites the existing sequence.
Add	Automatically changes the sequence name and adds the sequence.
Cancel	Only the sequence with the same name is not imported.
Cancel All	If multiple sequences are imported, all of them are not allowed.

About Comments

A comment is automatically given to a sequence that has been read from a file with any of the following formats: FASTA, GenBank Flat, EMBL, PIR, and former DNASIS.   For the comment-giving rules, refer to the description for the file.

To display a comment, click the (≡) button on the View Toolbar.   To hide the comment, click the button again.   It is also possible to edit comments directly.

If several sequences are read and displayed, the comment displayed corresponds to the target or activated sequence in sequence view pane. The name of the current target sequence is displayed in the Sequence View and the Map View. The current target sequence name is displayed on the Status Bar as well.

Target: "YO7510"; Position: 1 bp
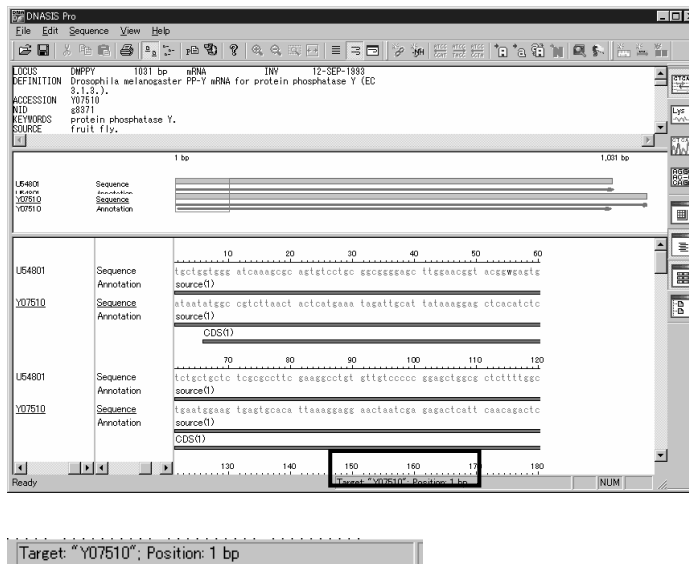
To switch the target sequence, click its sequence name or the sequence itself.

Upper Limit on the Number of Sequences

The maximum number of sequences that can be read in a single window is 100, including the number of newly created sequences.   The value is actually equal to the sum of the numbers of DNA sequences and amino acid sequences.

2.4 Editing Sequences (basic)

About the Insertion Pointer

If you click somewhere on a sequence, a vertical bar flashes, as shown in the figure, at the click point.   The bar is called the Insertion Pointer.   Any character entered from the keyboard is inserted at the Insertion Pointer.   The Insertion Pointer also serves as the starting point for keyboard operations, such as the process of deleting sequences.
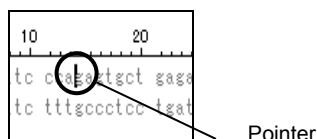
Pointer

Ways of Moving the Insertion Pointer

There are several ways of moving the Insertion Pointer.

The "cursor Up" key:	Moves the pointer to the same position one line above the sequence.
The "cursor Down" key:	Moves the pointer to the same position one line below the sequence.
The "cursor Left" key:	Moves the pointer back one character.
The "cursor Right" key:	Advances the pointer by one character.
The "Home" key:	Moves the pointer to the start of the sequence.
The "End" key:	Moves the pointer to the end of the sequence.

Inserting and Deleting Sequences

With the Insertion Pointer flashing, you can insert and delete sequences in the following procedures.

The "Character input" area:	Inserts characters that have been entered[1].
The Edit and Paste:	Inserts the content of the Clipboard[2].
The "Ctrl + V" key combination:	Inserts the content of the Clipboard[2].
The "Del" key:	Deletes the single character to the right of the Insertion pointer.
The "Back Space" key:	Deletes the single character to the left of the Insertion pointer.

[1]You cannot use any invalid characters as DNA sequences or amino acid sequences.   For the list of characters you can use, refer to "Characters You Can Use for DNA Sequences" and "Characters You Can Use for Amino Acid Sequences" in Section 2.2.

[2]Any invalid characters as DNA sequences or amino acid sequences will be removed.

Pasting from the Clipboard

Select Edit and then Paste or press the Ctrl + V key combination to paste the content of the Clipboard.   Note, however, that the operation may vary depending on the working conditions.

With the Insertion Pointer flashing:	The content of the Clipboard is inserted as a sequence.
With a range selected:	The selected range is replaced with the content of the Clipboard.

Any characters invalid as DNA sequences and amino acid sequences are automatically removed before the paste takes place.

Selecting the Range

If you drag part of a sequence using the mouse, the color of the dragged area changes.   This highlighted area is called a selected range.

The selected range lets you perform a variety of operations, including deleting, replacement, changes between uppercase and lowercase characters, interconnection with the result of analysis, and annotations.
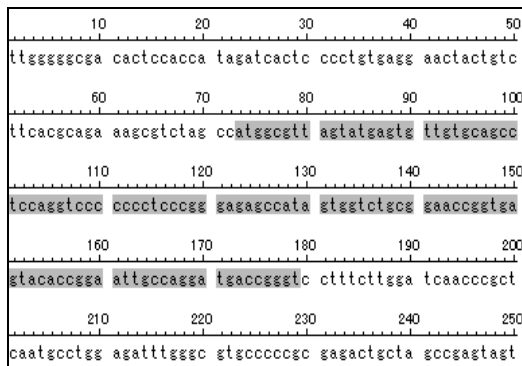
Ways of Selecting a Specific Range

There are several ways of making selected ranges.

Using the mouse:	Press the mouse's left button at the starting point and move the mouse to the ending point and release the button. You also can select more than one line at a time. Automatic scrolling starts if you move the mouse cursor outside the Sequence View.
Using the keyboard:	Move the Insertion Pointer to the starting point and, while holding down the Shift key, press the cursor move key.
Selecting a specified range:	Select Edit and then Select Range... and then enter a bp-measured value to specify the range you want to select.
Expanding a selected range:	While holding down the Shift key, click the mouse's left button or press the cursor move key. In contrast, however, it is impossible to reduce the selected range. In that case, first cancel the selected range, and then redo it.
Selecting an entire sequence:	Select Edit and then Select All.

Canceling the Selection

Click somewhere on a sequence or press the cursor move key.

Deleting the Selected Range

To delete a selected range, press the Del key or the Back Space key.   For more than one sequence, be sure that the sequence is handled as the target*: that is, its sequence name and analysis name are underlined.

Replacing the Selected Range

*Refer to "About the Target" in "2.9 Editing and Analyzing Multiple Sequences".

To replace a selected range, enter data from the keyboard or paste it from the Clipboard onto the selected range.   For more than one sequence, be sure that the sequence is handled as the target*: that is, it sequence name and analysis name are underlined.

Renaming Sequences

You can change the sequence name, which is found at the leftmost column of the Sequence View or Map View, using the following procedures.

1. In the Sequence View, click a sequence name you want to change.   The name then becomes the target and is now underlined.

*Refer to "About the Target" in "2.9 Editing and Analyzing Multiple Sequences".

2. Click the name again.

3. After a 0.5-second delay the outer frame is displayed, in which you can perform editing, as shown in the figure.
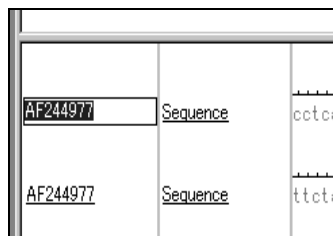
4. After editing, press Enter or click somewhere outside the frame.

Sequence names involve usable character and length limitations*.

Restrictions for Naming Sequences

Sequence names have limitations on the length and characters that can be used.

Characters that can not be used: < > ? * \ : / |

Length: Up to 128 characters

*Refer to "Restrictions for Naming Sequence" in "2.9 Editing and Analyzing Multiple Sequences".

2.5 Analyzing Sequences (basic)

Analyzing Sequences

1. From the Analysis Button view on the Main Window, find the analysis you want to perform.

2. Click the Analysis button, as shown in the figure.
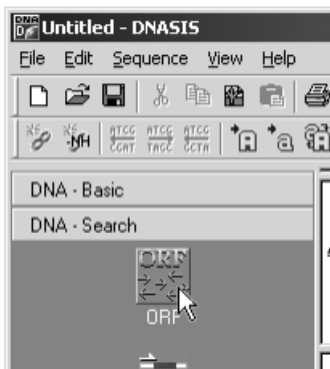
3. If you perform an analysis in which parameters can be set or edited by you, an Analysis dialog will appear before starting the analysis.

4. The analysis result will appear below the sequence or in a separate window.

Tabs classify the Analysis buttons*.

How to Display the Result of Analysis

There are two ways of displaying the result of analysis.

Inline view:	The result of analysis is displayed below a sequence so that the result is synchronized with the sequence. It appears in both the Sequence View and Map View. Types of analysis cover the GC content, ORF, translation, and restriction enzyme search.
Another window:	The result of analysis is not displayed in the Main window but in another window. Results of analysis cover the frequency of codon use and the phylogenic tree.
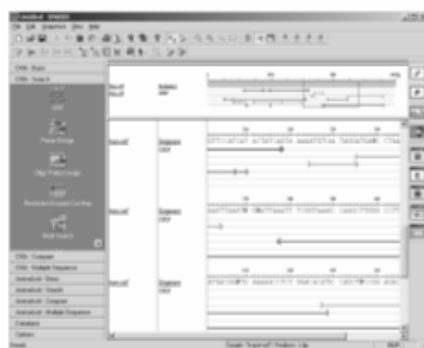
Inline Window   **Another Window**

The type of analysis automatically decides whether display will be in a new window or inline.

Customizing the Display of Analysis Results

*Refer to Chapter 3 "Details of Analysis".

You can customize the display of results of analysis; for example, changing the color and hiding some of the results.

Right-click the results of analysis to display the menu where you can perform operations.   For details, refer to the description of individual analysis results*.

Changing Analysis Parameters

Some items of analysis require parameter setting and may allow the default settings to be edited by you.   Examples

include database selection for homology search, enzyme type selection for restriction enzyme search, and codon tables

for translation. Use the following procedures to carry out such parameter-based analysis.

1. Click the relevant analysis button.

2. When the Analysis dialog box appears, click the Parameter button.

3. The parameter-setting dialog box appears (This is different depending on the type of analysis).

4. Set a parameter or parameters.

5. In the dialog box, click the OK button to close the dialog box.

Redoing Analysis

After, for example, editing sequences or changing parameters, you may want to redo the analysis.   In that case, use the

following procedures.

1. Click the Analysis button again.

2. When the Analysis dialog box appears, click the Execute button. The following dialog box appears.   In response,
   click the Overwrite button.
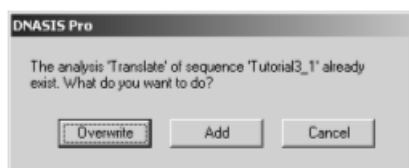
3. The result of analysis is overwritten.   If you click the Add button in step 2 instead, a new result of analysis is added to
   the display.

Deleting the Result of Analysis

You can delete the result of analysis in the following way.

1. Left-click the analysis name for a result of analysis you want to delete.   Once you left-click to select it, the analysis
   name is underlined, as shown in the figure.
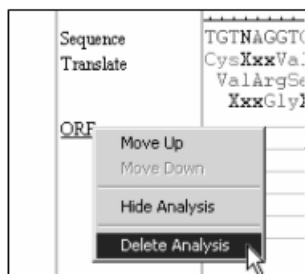
2. Right-click the analysis name.

3. When a menu appears, select Delete Analysis.

Hiding the Result of Analysis

You can hide the result of analysis temporarily.

1. Left-click the analysis name for a result of analysis you want to delete.   Once you left-click to select it, the analysis
   name is underlined, as shown in figure.
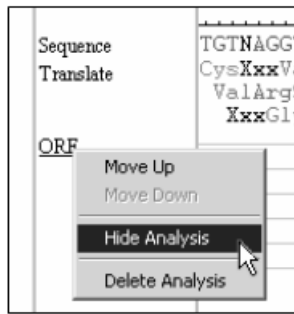
*Refer to "Redisplaying the Result of Analysis" in "2.5 Analyzing Sequences (basic)".

2. Right-click the analysis name.

3. When a menu appears, select Hide Analysis.

This redisplays the hidden result of analysis*.

Redisplaying the Result of Analysis

You can redisplay a hidden result of analysis.

1. Click the (▦) button on the View Toolbar; alternatively, you can click View and then Data List....

*Refer to "1.7 Data List Window".

2. The data list dialog box* appears.

3. Select the check box for a result of analysis you want to redisplay, so that a checkmark is placed in the box.

4. Press the OK button to close the dialog box.

2.6 Customizing Display of Sequences

The number of characters per line to display for sequences or the results of analysis can be selected from three choices:

No folding back characters; folding back characters according to the window width, and folding back characters

according to a specified width.

No Folding Back Characters
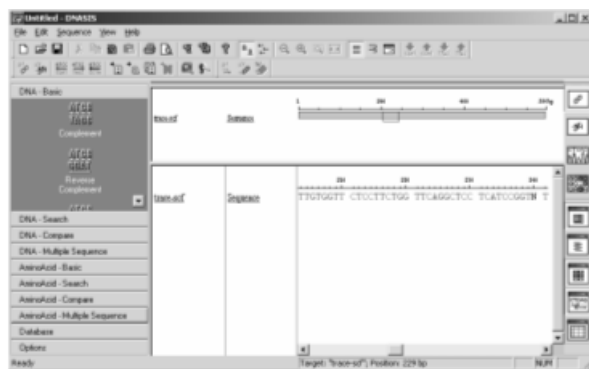
One line is used to display without folding back characters to a specified width.

1. Click the (≡) button on the toolbar.

2. Alternatively, select View and then Preferences....   In response to a dialog box that then appears, set Fold Sequence
   on the Folding/Ruler page to No Fold.

Using the horizontal scroll bar, you can scroll through the part extending past the viewable area of the window.

Folding Back Characters According to the Window Width

With this method, the characters to be displayed are folded back according to the width of the window.   Changing the

window size automatically changes the fold-back width accordingly.
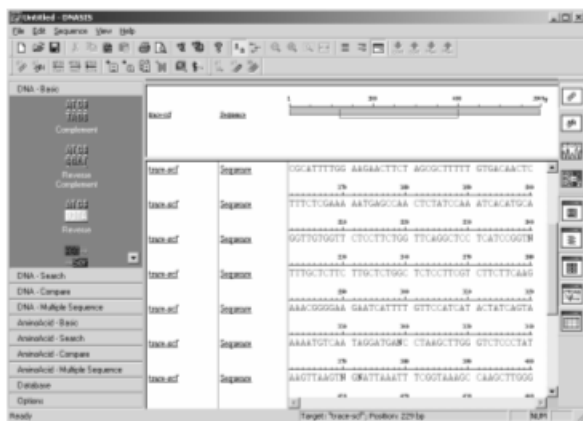
1. Click the (→|) button on the toolbar.

2. Alternatively, select View and then Preference... .   In response to a dialog box that then set Fold Sequence on the
   Folding/Ruler page to Fold by window width.

In the block-based display mode, the number of fold-back characters is changed according to a multiple of the block

length.   Otherwise, the number of fold-back characters is changed on a character basis.

Folding Back Characters According to a Specified Width

With this method, the characters to be displayed are folded back according to a specified number of characters.

1. Click the ( ) button on the toolbar.

2. Alternatively, select View and then Preference....   In response to a dialog box that then appears, set Fold Sequence on the Folding/Ruler page to Fold by every xx bp/aa.
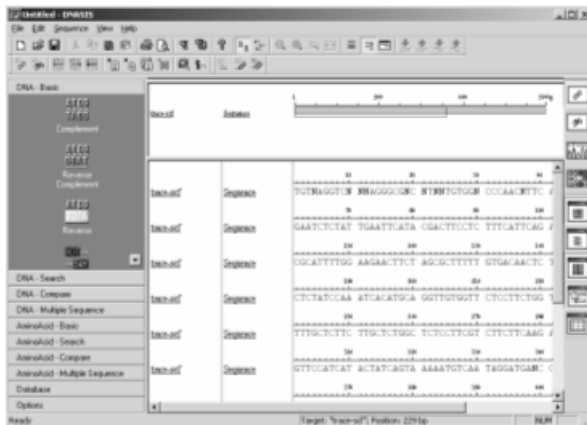
If block-based display mode is chosen in conjunction with the options to fold back characters according to a specified width, the value for the block length must divide the value of specified width without a remainder.

Inserting Spaces after a Specified Number of Characters (Block-Based Display Mode)

The "block-based display mode" makes it possible to insert a space into characters each time a specified number of characters is reached.

1. Select View and then Preference....   Alternatively, you can click the ( ) button on the toolbar.

2. In the dialog box that then appears, select the Folding/Ruler tab.
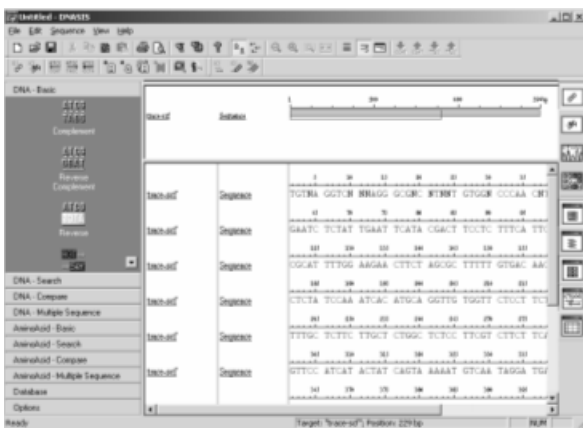
3. Place a checkmark in the Block Length check box.

4. Enter a value into the Block Length item to serve as the block length.   For the method of folding back characters according to a specified width (Fold by every xx bp/aa), you can specify only a value by which the line width can be divided without a remainder.

5. Press the OK button.

Hiding the Ruler

You can hide the ruler from the Sequence View.
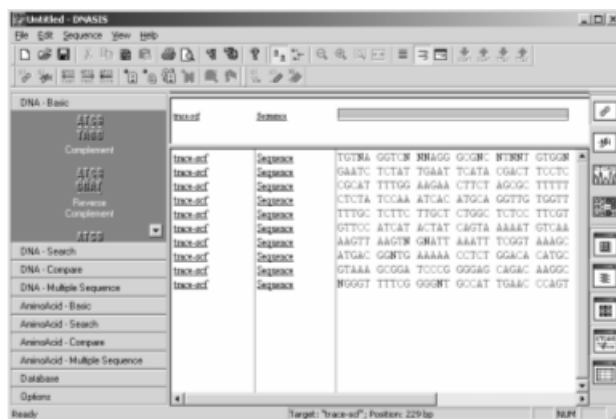
1. Select View and then Preference.... Alternatively, you can click the (⌨) button on the toolbar.

2. From the dialog box, select the Folding/Ruler tab.

3. Uncheck the Show Scale item.

4. Click the OK button.

Ways of Displaying the Ruler

There are several ways of displaying the ruler. Examples are the methods of using the scale line and bp indication.

1. Select View and then Preference.... Alternatively, you can click the (⌨) button on the toolbar.

2. From the dialog box, select the Folding/Ruler tab.

3. Perform the ruler setting.

Line & bp/aa:	Displays both the scale line and the bp indication above the sequence. For the alignment-based display, it shows the bp count for consensus sequences.
bp/aa	Displays only the bp indication above the sequence. For the alignment- based display, it shows the bp count for consensus sequences.

Line&bp/aa **bp/aa** **Line**
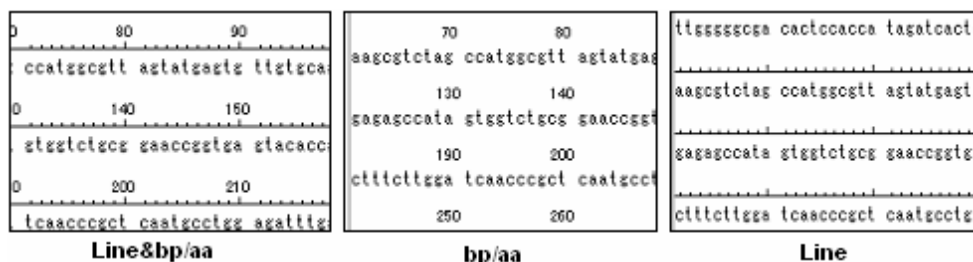
Line	Displays only the scale line above the sequence.	
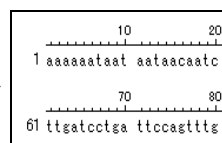
Show positions at the sequence head.

Assigns the bp indication to both the right and left ends of each line of a sequence. For alignments, the value is smaller by the gap.

4. Click the OK button.

Changing the Font for Sequences

You can change the font for sequences.

1. Select View and then Preference.... Alternatively, you can click the (⌨) button on the toolbar.

2. From the dialog box, select the Font tab.

3. Select Sequence in the combo box at the top.

4. Using the Setting button, set the font.

5. Click the OK button.

Select a font with equal width; otherwise, the display may crash.   Note that the color setting here is ignored.

Changing the Color of Sequences

You can change the color of sequences.

1. Select View and then Preference....   Alternatively, you can click the (![icon]) button on the toolbar.

2. From the dialog box, select the Sequence tab.

3. Perform the setting within the Sequence Color box.

Colorize sequence view:	When checked, this item displays sequences in the color mode. When unchecked, it provides a black display.
Other item:	Sets the color on a character basis. You can set the color of characters and the color of the background separately. DNA sequences and amino acid sequences are also set separately.

4. Click the OK button.

Displaying Pre-Edit Original Sequences

You can display pre-edit original sequences (those sequences available immediately after they are read from a file) at

the same time.

1. Select View and then Preference....   Alternatively, you can click the (![icon]) button on the toolbar.

2. From the dialog box that then appears, select the Sequence tab.

3. Place a checkmark in the Show original sequence check box.

4. Click the OK button.

Sequences are displayed in a two-row pattern: the top row for original sequences and the bottom row for sequences

being edited.

Displaying Complement Sequences

You can display the complement sequences of sequences being edited at the same time.

1. Select View and then Preference....   Alternatively, you can click the (![icon]) button on the toolbar.

2. From the dialog box, select the Sequence tab.

3. Place a checkmark in the Show complementary sequence check box.

4. Click the OK button.

5. Sequences are displayed in a two-row pattern: the top row for sequences being edited and the bottom row for
   complement sequences.   The content of the bottom row is automatically updated while it is synchronized with the
   process of editing the content of the top row.

AF165046	Sequence	ttggggcgc cactccacca tagaccactt ccctgtgagg aactac aaccccgccg gtgaggtggt atctggtgaa gggacactcc ttgatg
AF165046	Sequence	aagcgtctag ccatggcgtt agtatgagtg tcgtgcagcc tccagg ttcgcagatc ggtaccgcaa tcatactcac agcacgtcgg aggtcc
AF165046	Sequence	gagagccata gtggtctgcg gaaccggtga gtacaccgga attgcc ctctcggtat caccagacgc cttggccact catgtggcct taacgg
AF165046	Sequence	ctttcttgga tcaacccgct caatgcctgg agatttgggc gtgccc gaaagaacct agttgggcga gttacggacc tctaaacccg cacggg
AF165046	Sequence	gccgagtagt gttgggtcgc gaaaggcctt gtggtactgc ctgata cggctcatca caacccagcg ctttccggaa caccatgacg gactat
AF165046	Sequence	gccccgggag gtctcgtaga ccgtgcatca tgagcacaaa tcctaa cggggccctc cagagcatct ggcacgtagt actcgtgttt aggatt

2.7 Editing Sequences (advanced)

Selecting Sequence Ranges

You can select a number of non-continuous ranges of a sequence at the same time.

1. Select the first range*.

2. While pressing the Ctrl key, drag the mouse and select another range.

3. These selected ranges are highlighted: the first one looks orange; the last one looks pink.

4. Repeat steps 2 to 3 above.

```
.......... .......... .......... .........
caccaaccgc cgcccacagg acgtcaagtt cccgggc

.......... .......... .......... .........
ttacctgttg ccgcgcaggg gccccaagtt gggtgtg

.......... .......... .......... .........
gcggt gcaa cctc tggaa ggc acaacc ta cccc

.......... .......... .......... .........
ggcctgggcc ggtacccctg gcccctctat ggcaatg
```
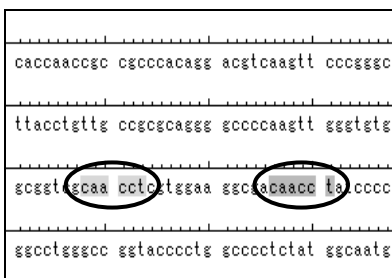
It is also possible to select two overlapping ranges.   The overlap does not have a special meaning when it comes to usual commands regarding the selected range.   For annotations, however, separate annotation entries* are created from the viewpoint of the overlap.

Converting Uppercase and Lowercase Characters

You can convert between the uppercase and lowercase characters of sequences.

Selecting Sequence and then Upper Case or clicking the ( ) button on the toolbar: Conversion from lowercase to uppercase characters

Selecting Sequence and then Lower Case or clicking the ( ) button on the toolbar: Conversion from uppercase to lowercase characters

Selecting Sequence and then Exchange Case or clicking the ( ) button on the toolbar: Conversion from lowercase to uppercase characters and vice versa

What is to be converted is different depending on the working condition -- whether or not there is a selected portion.

Yes:	Converts only the selected portion of the sequence. If there are several selected portions, all of them are converted.
No:	Converts the entire of the sequence.

Masking Sequences

You can mask selected portions of sequences.   The masked portions are replaced with N for DNA sequences and with X for amino acid sequences.   Masking makes it possible to skip the selected portions to be analyzed.

1. Select Sequence and then Mask or click the ( ) button, or the ( ) button for amino acid.

If there are several selected ranges, all of them are masked.

Converting into Complement Sequences, Reverse Complement Sequences, and Reverse Sequences

Sequences being edited are converted into the following: complement sequences, reverse complement sequences, and reverse sequences.

Selecting Sequence and then Complement or clicking the ( ) button on the toolbar: Conversion into complement sequences

Selecting Sequence and then Reverse or clicking the () button on the toolbar: Conversion into reverse sequences

Selecting Sequence and then Reverse Complement or clicking the () button on the toolbar: Conversion into the reverse sequences of complement sequences.

The entire of a sequence being edited undergoes the process of conversion.

Returning to the Pre-Edit Original Sequences

This process is intended to cancel all changes made on a sequence being edited so that the sequence will be returned to the original state immediately after it was read from a file.

1. Select Sequence and then Revert.

2. In the confirmation box, click the OK button.

2.8 Analyzing Sequences (advanced)

Displaying Results of Analysis Side by Side

The results of analysis from different analysis buttons are automatically displayed in a vertical format.   Usually, the result of analysis of the same type is overwritten.   However, you may want to avoid overwriting during such analysis; for example, when you have changed parameters or edited sequences.   In that case, proceed as follows.

1. Perform the first analysis.
2. Perform operations such as changing parameters or editing sequences.
3. Repeat analysis.
4. In response to a message saying that "The analysis 'xxx' of sequence 'xxxxx' already exist. What do you want to do?", click the Add button.
5. Preferably, you should change the analysis name*.
6. Repeat steps 2 to 5, as necessary.

*Refer to "Changing Analysis Names" in "2.8 Analyzing Sequences (advanced)".
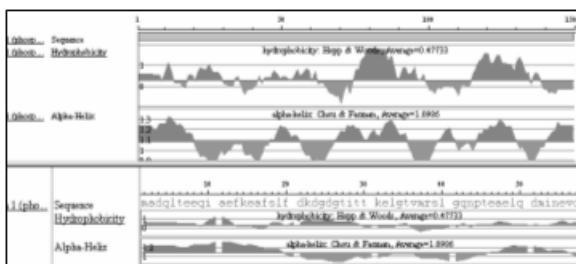
As a result, you can display as many analysis results as you like side by side.

Interlocking the Range of Selection among Results of Analysis

You can interlock the selected ranges of multiple analysis results.

With several analysis results for the same sequence displayed, providing the analysis results or the sequence with range selection causes the selected range to be interlocked automatically.

Such an automatic interlock occurs even when there are several selected ranges.   This makes it relatively easy to compare the locations of functional parts of the sequences.
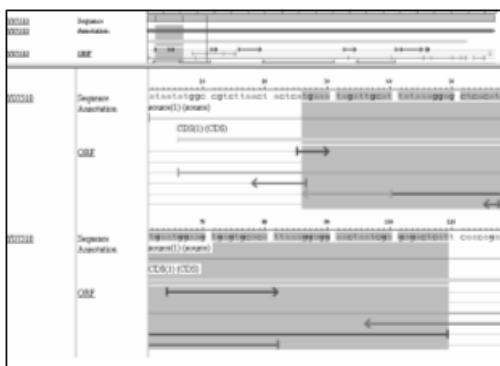
Creating Analysis Buttons Having Different Parameter Setting

Changing parameters each time analysis is performed can be tedious. Frequent changing of parameters could be needed in such situations as when selecting a database for homology search, selecting an enzyme type for restriction site search,

and providing a codon table for translation.   You can solve this problem if you duplicate analysis buttons and associate

each with set of desired parameter settings. Right-click an analysis button you want to duplicate.

1. From the menu, select Duplicate.
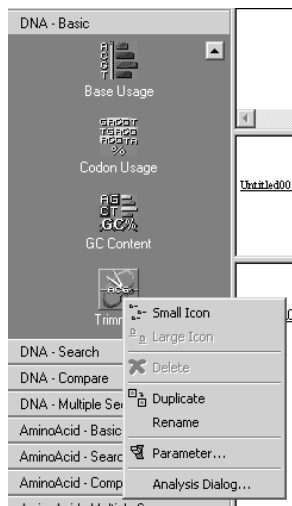
2. An icon is duplicated just below the button, and the button name is ready for editing.   Change the button name to
   something else.
3. Click the duplicated icon and an Analysis dialog will appear. Then click the Parameter button.
4. When the parameter-setting dialog box appears, change all parameters.
5. Click the OK button to close the dialog box.

Changing Analysis Names in display

In column 2 of the Sequence View, the analysis name is displayed to the left of each result of analysis.   By default, the

name of the analysis button used to perform the analysis is set here.   You can manually change the displayed analysis

name in the following way:

1. In the Sequence View, click an analysis name you want to change.   The particular name then becomes the target,
   which is now underlined.
2. Click the name again.
3. After a 0.5-second delay, the outer frame is displayed, in which you can perform editing.
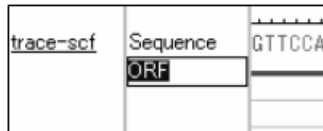
4. After editing, press Enter or click somewhere outside the frame.

Renaming Analysis Buttons

You can change the analysis name associated with the button.   Right-click an analysis button you want to duplicate.

1. From the menu, select Rename.
2. Since the button name is now ready for editing, as shown in the figure, change the button name.
3. After editing, press Enter.

Deleting Analysis Buttons

You can delete user-duplicated analysis buttons.

1. Right-click an analysis button you want to delete.

2. From the menu, select Delete.

Note that you are only allowed to delete those analysis buttons that you created by duplication, not the original ones.

Changing the Order of Analysis Display

You can change the order in which categories of analysis are displayed to provide an easier-to-read, well-organized

result.

1. Click the name of analysis whose order you want to change. This will make the analysis name ready to be selected.

2. Click the 🔼 or 🔽 button on the toolbar.   Alternatively, you can right-click the analysis name; in response to a

   pop-up menu that then appears, select Move Up or Move Down.

3. Selecting Move Up or Move Down thus changes the order of analysis results.

Repositioning Analysis Buttons

You can change the order of analysis buttons by dragging and dropping a button or buttons.

2.9 Editing and Analyzing Multiple Sequences

Creating New Sequences

You can create or add new sequences in a window that displays existing sequences.   Such new sequences are added at the end of the list of existing sequences. You, however, can change the order of these sequences*.

1. Make sure that the target sequence has not undergone range selection.

2. Select Sequence and then New DNA or click the ⌖ button on the toolbar.

3. A sequence whose name begins with Untitled001 is added to the end of the sequence list.

4. At the Insertion Pointer, which is flashing at the start of the sequence, enter an appropriate sequence from the keyboard, as shown in the figure.

Creating a New Sequence from Part of an Existing Sequence

You can extract any range of any sequence so it can serve as another sequence.

1. Select any range of any sequence, as shown in the figure.

2. Select Sequence and then New DNA or click the ⌖ button on the toolbar.

3. A sequence whose name begins with Untitled001 is added to the end of the sequence list.   The sequence corresponding to the range selected in step 1 is copied, as shown below.

Creating New Sequences by Linking Noncontinuous Ranges of an Existing Sequence

You can join several noncontinuous ranges of any sequence and extract them so that they can serve as another sequence. This function is convenient, for example, when you want to select all ranges of the exon part before creating a new sequence by joining them.

*Refer to "Selecting Ranges" in "2.7 Editing Sequences (advanced)".

1. Select any number of ranges of any sequence*, as shown in the figure.

2. Select Sequence and then New DNA or press the button on the toolbar.
3. A sequence whose name begins with Untitled001 is added to the end of the sequence list.   The sequence corresponding to the ranges selected in step 1 is duplicated (by joining them from left to right), as shown in the figure.

*Refer to "About the Target" in "2.9 Editing and Analyzing Multiple Sequences".

4. If different sequence range is found duplicated or nothing is found duplicated, make sure that the regions of the sequence which are to be duplicated have actually been range-selected as the target*.

Duplicating the Sequences Entirely

You can duplicate the sequence entirely.

*Refer to "About the Target" in "2.9 Editing and Analyzing Multiple Sequences".

1. Select a sequence you want to duplicate as the target*.
2. Select Sequence and then Duplicate.
3. All of the sequence selected in step 1 is duplicated under the "Original sequence name" + "Copy" name at the end of the sequence list.

The original sequence and gap information are also duplicated, although the trace data or annotation is not duplicated.

Reading New Sequences from a File

There are two ways of reading sequences from a file and adding them to a window.

1. Select File and then Open… or File and then Import Sequence… before selecting files you want to read from a list of files.   In this case, you can also select several files at the same time.
2. Using Windows Explorer, select a file you want to read, and drag and drop it to the DNASIS window.   It is also possible to drop in several files at the same time.

Renaming Sequences

You can change the name of any sequence.   For details, refer to "Renaming Sequences" in "2.4 Editing Sequences (basic)".

Restrictions for Naming Sequences

Sequence names have restrictions concerning their length and font type.

Hiding Sequences

You can temporarily hide any sequence and the analysis result for the sequence.

1. Right-click the name of a sequence you want to hide.

2. From the pop-up menu, select Hide, as shown in the figure.
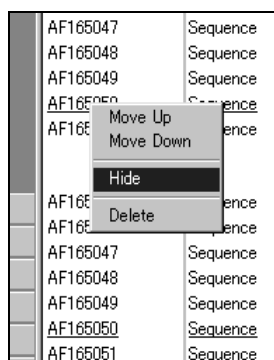
3. This action hides sequence and its analysis result.

To redisplay the hidden sequence and its analysis result, click the ⊞ button on the View Toolbar and respond to a dialog box that appears*.

*Refer to "1.7 Data List Window".

Deleting Sequences

You can delete any sequence and its analysis result.

1. Right-click the name of a sequence you want to delete.

2. From the pop-up menu, select Delete, as shown in the figure.
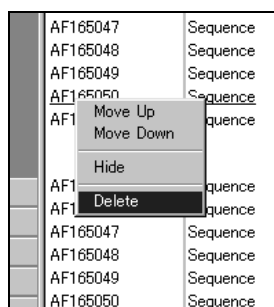
3. Selecting Delete will delete the sequence and its analysis result.

Once any sequence or its analysis has been deleted, you cannot restore it.

Changing the Order of Sequence Display

You can change the order in which sequences are displayed to provide an easier-to-read, well-organized result.

1. Click the name of sequence whose order you want to change, so that the sequence name is ready to be selected.

2. Click the ⬆ or ⬇ button on the toolbar.   Alternatively, you can right-click the sequence name. From the menu, select Move Up or Move Down.
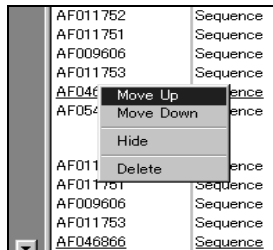
3. This selection changes the sequence and its analysis result.

Even in the alignment display mode, it is possible to change the display order.

About the Target

Usually, a single sequence is the target of analysis, except for two analysis groups: the DNA - multiple-sequence and the amino acid - multiple-sequence.   If you click either of those analysis buttons, some sequences become the analysis target.   These sequences called the "target" have their sequence names underlined.   The sequence name for the current target also appears on the status bar located at the bottom of the window.   You can switch the target by clicking the sequence name.

Selecting Sequences as the Target of Editing

Clicking a sequence you want to edit causes the Insertion Pointer to appear there so that you can edit it.

Selecting Sequences as the Target of Analysis

If you left-click the name of a sequence to analyze, the sequence is set as the target and it is underlined.   The sequence name for the current target also appears on the status bar located at the bottom of the window.

Analyzing Multiple Sequences at Once

Multiple sequences are required for two types of analysis groups: the DNA - multiple-sequence and the amino acid - multiple-sequence.   Therefore, the target of analysis covers not only the target sequence but also all the sequences that are currently displayed.   To remove it from being analyzed, hide the current sequence temporarily*.

*Refer to "Hiding Sequences" in "2.9 Editing and Analyzing Multiple Sequences".

Usually, a single sequence is the target of analysis, except for two types of analysis groups: the DNA - multiple-sequence and the amino acid - multiple-sequence.   If you click the analysis button while holding down the Ctrl key, the analysis covers all the sequences that are currently displayed.
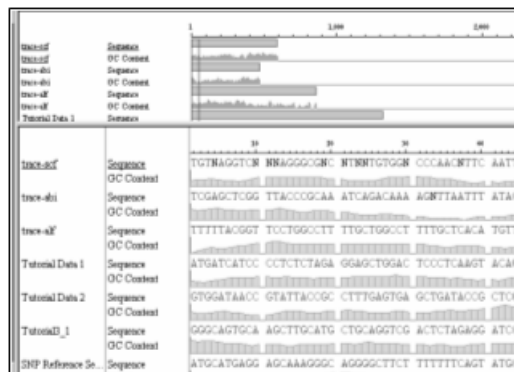
2.10 Searching for Sequence strings

Searching for Sequence Strings

Using a character string, this function searches for a string in the sequence being edited.

1. Select Sequence and then Find… or click the ![icon] button on the toolbar.

2. The following dialog box appears.

3. Enter a character string you want to search for in the Find What field.

4. Click the Find button.

5. If a match occurs, the window automatically scrolls to the range of the match.

The search process is case-sensitive; uppercase and lowercase characters are distinguished.   Search starts at the point

where the Insertion Pointer is currently located or at the point following the selected range.

Jumping to the Next Match

To jump to the next match, select Sequence and then Find Again or press the F3 key.   To go back to the previous match,

press both Shift and the F3 key at the same time.

Selecting All Matches at Once

1. Select Sequence and then Find… or click the ![icon] button on the toolbar.

2. The following dialog box appears.

3. Enter a character string you want to search for in the Find What field.

4. Click the Find All button.

5. The range of all the matches found is selected.   They all are colored orange, except for the last one, which is pink.

6. To jump to the next match, select the Sequence and then Find Again or press the F3 key.

The search process is case-sensitive; uppercase and lowercase characters are distinguished.   Matches are colored in the

Map View so that you can, at a glance, see the distribution of matches over the entire sequence.

Selecting Sequences as the Target of Search

*Refer to "About the Target" in "2.9 Editing and Analyzing Multiple Sequences".

Normally a search handles the sequence that is currently selected as the target.   If you want to search within another

sequence, you must first set that sequence as the target*.

Searching Multiple Sequences

You can select multiple sequences within the same project as the target of a search for a sequence string.

1. Select Sequence and then Find… or click the ![icon] button on the toolbar.

2. The following dialog box appears.

3. Enter a character string you want to search for in the Find What field.

4. Select All Sequences in the Find Range field.

5. Click the Find button.

If a match occurs, the window automatically scrolls to the range of the match.   To jump to the next match, select Sequence and then Find Again or press the F3 key.   The "next match" actually refers to the sequence reached in the process of searching the list of sequences currently displayed on the window in the following order: from left to right and from top to bottom.

It is also possible to search for, at once, all matches from all the sequences that are currently displayed.

2.11 Annotations

About the Annotation

You can add information to a specified range of a sequence.   For example, the GenBank format has the FEATURES

table, which contains pieces of information about a sequence, such as the CDS region and promoter region.   DNASIS

MAX is designed to extract information from tables and display it graphically.

A description or other information to be associated with a sequence or a region of a sequence is called an annotation.

The individual pieces of information added to describe the annotations in detail are called "annotation entries".

Creating New Annotations

1. Select a sequence for which you want to create a new annotation.

2. Click the ⬚ button on the toolbar.

Creating Annotation Entries

1. Select a sequence for which you want to create a new annotation entry.

2. If there are more than one annotation click the annotation with which you want to associate the annotation entry.

3. Click ⬚ on the Toolbar. The Annotation Setting dialog appears.   The annotation entry is associated with the
   annotation you clicked . If there was no designated annotation, a new annotation is automatically created.

4. Enter the Annotation Name and Annotation Kind.

5. Enter the value to create an annotation in Annotation Range.

6. Specify the direction of annotation entries in the Direction field.

7. Click the OK button.

Assigning Annotation Entries to the Range of Selection

1. Select the range of an appropriate sequence in the Sequence View.

2. Click [icon] on the Toolbar. The Annotation Setting dialog appears. The value of selected range will be automatically set in Annotation Range.

3. Enter the Annotation Name and Annotation Kind.

4. Enter the Orient value.

5. Click the OK button.

Assigning Annotation Entries to Multiple Ranges of Selection at Once

*Refer to "Selecting Ranges" in "2.7 Editing Sequences (advanced)".

1. In the Sequence View, select several ranges of a sequence*.

2. Click the [icon] button on the toolbar.   Each annotation entry is named Unknown.

Editing Annotation Entries

1. In the Sequence View, select an annotation entry you want to edit.

2. Double-click the annotation entry.   Alternatively, you can right-click the annotation entry and select the Edit Annotation... menu item.   The Annotation dialog box appears, as shown in the figure.   The Annotation Setting dialog appears.
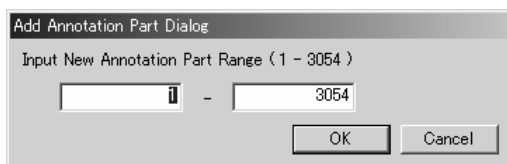
3. To add a part, click Add under Part Range to display the Add Annotation Part Dialog. Specify the range for the part, then click OK to return to the Annotation Setting dialog.

4. To add comments, click Add under Comment to display the Add Annotation Comment Dialog. Enter the comment key and value, then click OK to return to the Annotation Setting dialog.
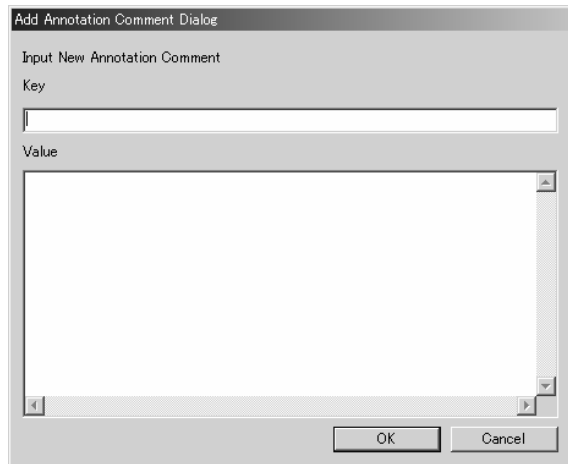
5. Edit in Annotation Setting dialog, then click OK to store the modification.

Deleting Annotation Entries

1. Select an annotation entry you want to delete.

2. Right-click the annotation entry and select the Delete Annotation menu item.

3. This selects the annotation entry.

Deleting Annotations

1. From the analysis name of column 2 in the Sequence View, select an annotation you want to delete.

2. Right-click the selected analysis name and select Delete Analysis in the menu that appears.

3. This deletes the annotation.

Creating Multiple Annotations

You can store annotations after dividing them into groups.   Suppose, for example, you want to add CDS information

and SNP information as annotations to genome sequences.   In this case, you can create annotations by dividing them

into those for CDS information and those for SNP information.

1. Select a sequence to which you want to add an annotation.

2. Click the [icon] button on the toolbar to create the first annotation.   The analysis name for the annotation is called Annotation.

3. Similarly, click the [icon] button on the toolbar to create the second annotation.   The analysis name for the resulting annotation is called Annotation 1.

4. Using its analysis name, select an annotation and click [icon] on the toolbar, so that an annotation entry is added to the specified annotation.

2.12 Printing

Printing the Map View

After clicking in any blank part in the Map View, select File and then Print... or click the button on the toolbar.

Printing the Sequence View

After clicking any blank part in the Map View, select File and then Print... or click the button on the toolbar.

Printing Only the Current Range of Display

1. Determine any part you want to print by using the layout view.

2. Select File and then Print Page... .

2.13 Projects

About the Project

The term "project" refers to a collection of the sequences that have been opened in a single editor, along with their

analysis results.   DNASIS lets you store sequences on a project basis.   Such a project is given a .dnasis extension.

Saving Projects

1. Select File and then Save Project or click the 🖫 button on the toolbar.   The Save As dialog box then appears.

2. Specify the storage location and file name before clicking the Save button.

Opening Projects

1. Select File and then Open or click the 📂 button on the toolbar.   The Open files dialog box then appears.

2. Specify the location for a project you want to open and the file name before clicking the Open button.

2.14 Waveform Display Mode

Entering Waveform Files

You can display a waveform file (also called trace file or chromatogram) that the DNA auto sequencer produces.   The waveform file can be read if its format is ABI or SCF.

1. After selecting File and then Open, select a waveform file you want to read.   It is also possible to select more than one file at the same time.

2. The sequences stored in the selected waveform file are then read and they are shown in the DNA display mode.

3. To display a waveform, click the ⸽⸽⸽ button on the View Toolbar, as shown in the figure.
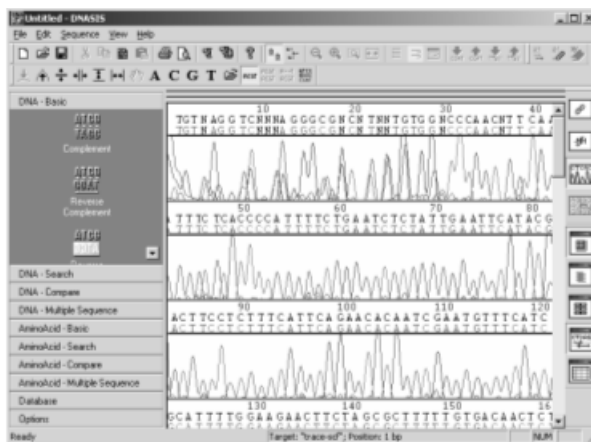
In the two-row format, the bottom row shows the original sequence that has been read from a waveform file.   In contrast, the top row shows a user-editable sequence although it is identical to the original sequence under the initial setting.

See the window below, which shows multiple items of data that have been displayed at the same time.   Since no fold-back display takes place, you need to scroll through them by means of the horizontal scroll bar.   Above the base sequence is displayed its sequence name.   The background for the sequence name is blue, which means that it has been selected as the target.   Therefore, it can be executed from the menu or toolbar button.   You can change targets by clicking somewhere on the trace data.
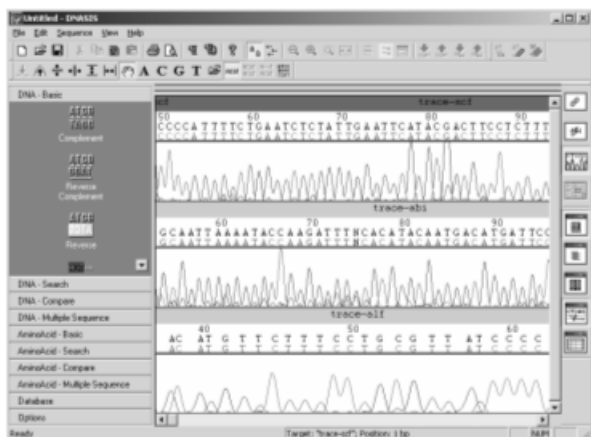
Switching between Waveform and Sequence Displays

You cannot analyze sequences in the waveform display mode.   In that case, you need to switch to the DNA display mode.   To display a sequence, click the ![icon] button on the View Toolbar.   To display a waveform, click the ![icon] button on the View Toolbar.   The method of a range-selecting sequence is convenient because the selected range interlocks two modes: the DNA display and the waveform display.

Selecting Waveforms to Be Displayed

You can display only specified waveform when there are several waveforms that have been read.

1. Click the ![icon] button on the View Toolbar to display the dialog box showing the list of analysis results.
2. Look for and select a line in which the Data Name field shows the data name and the Analysis Name fields gives Trace.
3. To display the data, click the Show button.   To hide it, click the Hide button.
4. Press the OK button.

Double-clicking the header for the Analysis Name filed in the list of analysis results causes the results to be sorted according to the order of the analysis names.   This function can be conveniently used when you want to select a line. Click while the Shift key is held down to select a range.   Click while the Ctrl key is held down to select more than one line.

Displaying Reverse Complement Sequences

You can display the reverse complement sequence of trace data.

1. Click the ![icon] button on the Waveform toolbar.

The waveform displayed is in reverse time order, resulting in a reverse complement sequence.   Switching to the DNA display mode under this condition will retain the status of the reverse complement sequence.   If there are several waveforms that have been displayed, click a target waveform that you want to display a complement sequence.

Editing Sequences While Viewing Their Waveforms

You can delete or replace bases while viewing their waveforms.

To delete bases, select a range you want to delete and press the Del key.

To replace a single base, select the base you want to replace and key in the new base.   Selecting 2bp or more will cause the replacement of the base to fail.

Use the following procedures to insert bases.

1. Viewing the waveform makes a range selection of 2bp of the base before and after the point into which you want to insert a base.
2. Click the ![icon] button on the View Toolbar to switch to the DNA display mode.
3. To use the Insertion Pointer, click somewhere in the highlighted 2bp range, which was selected in step 1.
4. Enter a base from the keyboard.
5. Select the range of the base entered in step 4.   (The range selection helps you recognize the region when you switch to the waveform display mode.)
6. Click the ![icon] button on the View Toolbar to switch to the waveform display mode.

Returning to the Original Condition when Editing

You can cancel the entire process of editing a sequence and return to the original sequence.

1. Click the ![icon] button on the View Toolbar to switch to the DNA display mode.

2. If there is more than one sequence, click the sequences to set them as the target.

3. Select Sequence and then Revert.

4. When a confirmation dialog box appears, click the OK button.

5. Click the ![icon] button on the View Toolbar to switch to the waveform display mode.

Hiding Specific Lanes

You can display or hide the A, C, G, and T waveforms.   If you click the $\mathbf{A}$, $\mathbf{C}$, $\mathbf{G}$, $\mathbf{T}$ or button on the

Waveform toolbar, the corresponding waveform is hidden and the base sequence italicized.   The toggle button allows

you to switch between the display and hide modes each time it is clicked.

Expanding and Shrinking Displayed Waveforms

You can change the vertical and horizontal scales of the waveform display area.

![icon] □Reduces the display area vertically.

![icon] □Expands the display area vertically.

![icon] □Reduces the display area horizontally.

![icon] □Expands the display area horizontally.

Changing the Color of Waveforms

You can change the colors of waveforms and bases.

1. Select View and then Preferences… and click the ![icon] button on the toolbar.

2. This displays the Preferences dialog box, in which you should click the Sequence tab.

3. Enter a value for the Sequence Color field for each base type.

4. Click the OK button to close the dialog box.

Making Alignments with Reference Sequences

Use of this function requires you to obtain a separate multiple alignment option.

You can calculate and display an alignment with respect to the reference sequence.   The feature of highlighting

non-matched sequences is extremely helpful in detecting SNPs.

1. Read in the trace data as a candidate of the target to display.

2. If more than one sequence is displayed, click the one you want to set as the target.

3. Click the ![icon] button on the Waveform toolbar and select a sequence file you want to use as the reference.   You can
   only specify a FASTA file here.   Click the Open button to close the dialog box.

4. The reference sequence is displayed at the top of the sequence list.

5. If you click the ![icon] button on the Waveform toolbar, the alignment is calculated and displayed.   The background of
   a non-matched sequence becomes blue.

6. To stop the alignment display, click the ![icon] button on the Waveform toolbar.

Scrolling through Multiple Waveforms Horizontally and Separately

You can scroll horizontally through each of the waveforms displayed.   This function lets you align different waveforms

at a specific bp position.

1. Read in several waveform files to display at the same time.

2. Click the ![icon] button on the Waveform toolbar, when the mouse cursor changes its shape to ![icon] .

3. Drag a waveform being processed.

4. Click the ![icon] button on the Waveform toolbar again to return to the normal mode.

Copying Trace Data

You can copy trace data into the Clipboard after converting it to numeric values or graphics.   This function is very helpful for report making because it lets you copy only a specified range in the form of graphics. This function also allows high-resolution printing of the copied graphics.

Use the following procedures to copy numeric data.

1. Drag a waveform to select a range, as shown in the figure.
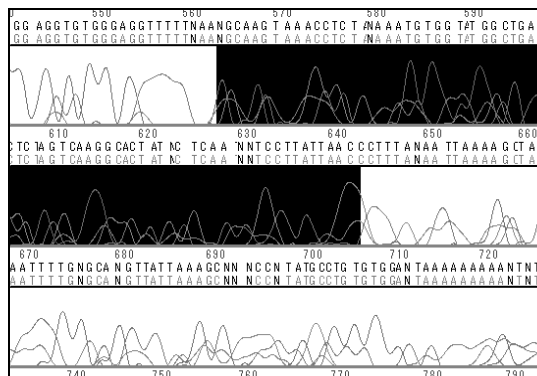
2. Right-click to display the pop-up menu.

3. Select "Copy Trace Value".

4. Paste the copy into another application such as MS-Excel, as shown in the figure.
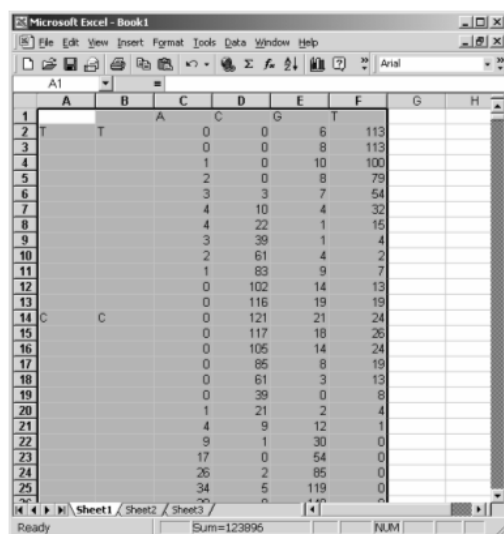
Use the following procedures to copy the graphics.

1. Drag a waveform to select a range.

2. Right-click to display the pop-up menu.

3. Select "Copy".

4. Paste the copy into another application such as MS-Word, as shown in the figure.
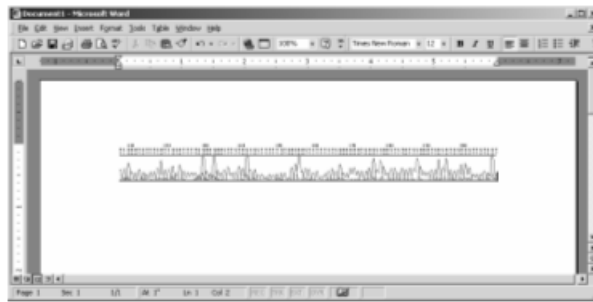
2.15 Saving Sequences as Text Files

You can output the sequences currently displayed in the window as a file with the FASTA format.

1. In the Sequence View, select a sequence you want to export.

2. Select Export... from File in the menu.
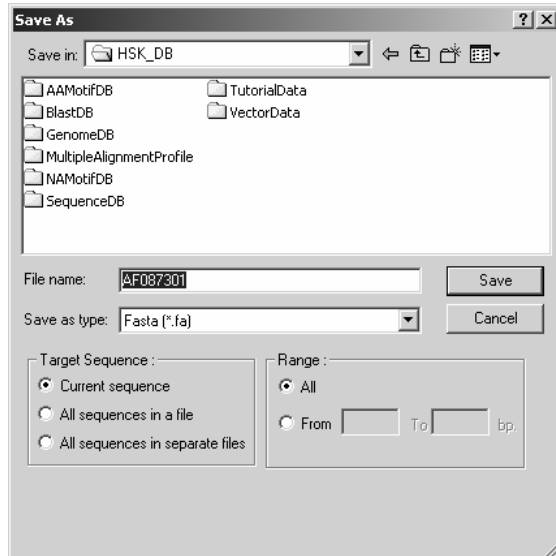
3. Enter a file name and click the Save button.

2.16 Copying Images

You can copy all the graphics displayed in the Map View or Sequence View into the Clipboard.   Since these graphics are actually copied as vector data, you can paste them to another application such as MS-Word and produce a high-resolution printout.

1. Click any blank part in the Map View or Sequence View to switch to the active mode.

2. Select a range you want to copy using such operations as expanding, shrinking, and scrolling.

3. Select Copy and then Copy Image or click the 🔳 button on the tool bar.

4. Switch to another application such as MS-Word and paste the copy.

Note: Copy pasting requires you to specify the following options: "Paste after Selecting Format - Graphics (Extended Metafile)".

2.17 Terminating DNASIS

1. From the File menu, select Exit.

Chapter 3 Analysis Functions

3.1 List of Analysis Functions

DNASIS MAX supports the following analysis functions.

Analysis Category	Analysis Button Name
DNA - Basic	Complement Sequence
	Reverse Complement Sequence
	Reverse Sequence
	Translation
	Base Content
	Codon Usage
	GC Content
	Vector and Low-Quality End Trimming
DNA - Search	ORF
	Primer Design
	Oligo-Probe Design
	Restriction Site Search
	Motif Search
	Searching a Motif Pattern
	Mutation Site Search
	Hairpin Loop Search
	Stacking Site Search
	Tandem Repeat Search
DNA - Comparison	BLAST Search
	BLAST Search (Protein DB)
	BLAST Search (Translation DB)
	One-to-One BLAST Search
	Internet BLAST Search
	Internet BLAST Search (Protein DB)
	Internet BLAST Search (Translation DB)
	Smith-Waterman Search
	One-to-One Smith-Waterman Search
	BLAST Search and Extraction
	Clustering
DNA - Multiple Sequence	Multiple Alignment
	Phylogenic Tree
	Multiple Alignment Tree View
	Creating Multiple Alignment Profiles
	Phylogenic Tree (Using Profiles)
	Sequence Assemble
Amino Acid - Basic	Amino Acid Content
	Isoelectric Points
	Hydrophilicity, Hydrophobicity, and Secondary Structure
Amino Acid - Search	Motif Search
	Common Motif Search
	Proteolytic Site Search
Amino Acid - Comparison	BLAST Search
	BLAST Search (Translation DB)
	One-to-One BLAST Search
	Internet BLAST Search
	Internet BLAST Search (Translation DB)
	Smith-Waterman Search
	One-to-One Smith-Waterman Search
Amino Acid - Multiple Sequence	Multiple Alignment
	Phylogenic Tree

| Creating Multiple Alignment Profiles |
| Phylogenic Tree (Using Profiles) |
| NCBI Entrez Search |

3.2 Complement Sequence

This function converts DNA sequences into complement sequences and then adds them as new sequences.

Explanation of the Result Window
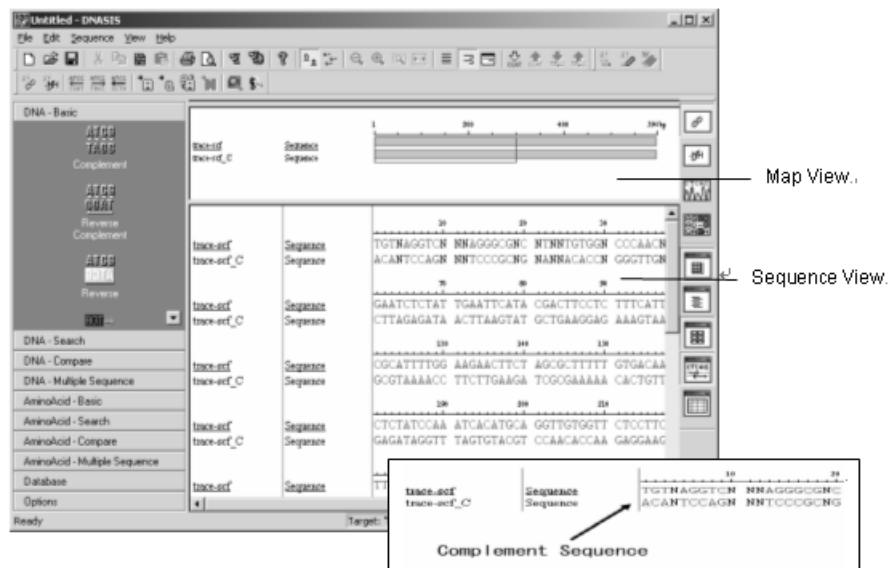
Sequence View

The complement sequence is displayed below the specified sequence.　The sequence name consists of the original sequence name followed by "_C".

 Example:

   If the sequence to be analyzed is ACTTGAGAT, clicking the Complement Sequence button converts it to TGAACTCTA.

3.3 Reverse Complement Sequence

This function converts DNA sequences into complement sequences, converts them into reverse sequences, and then adds the converted complement sequences as new sequences.

Explanation of the Result Window

Sequence View

The reverse complement sequence is displayed below the specified sequence.   The sequence name consists of the original sequence name followed by "_RC".

 Example:

> If the sequence to be analyzed is ACTTGAGAT, clicking the Reverse Complement Sequence button
>
> converts it to ATCTCAAGT.

3.4 Reverse Sequence

This function converts DNA sequences into reverse sequences and then adds them as new sequences.

Explanation of the Result Window
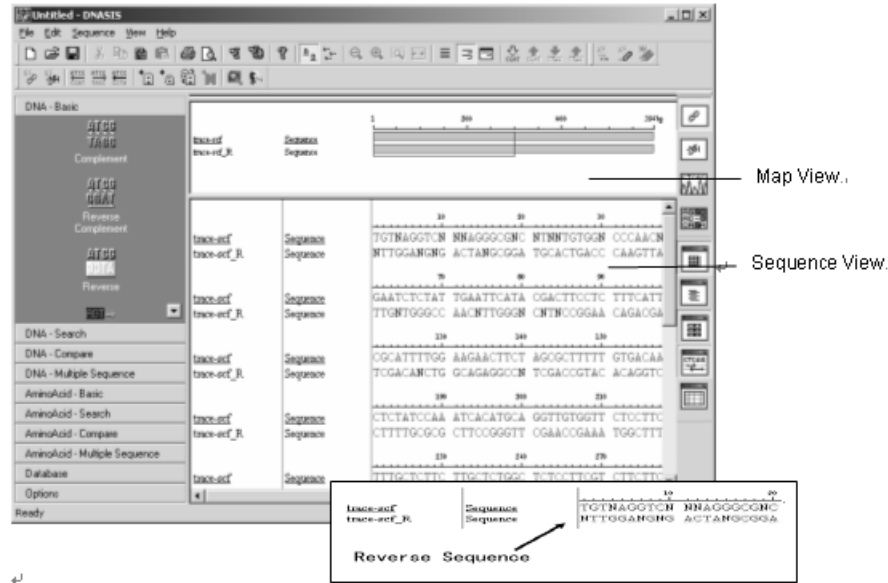
Sequence View

The reverse sequence is displayed below the specified sequence.   The sequence name consists of the original sequence name followed by "_R".

 Example:

If the sequence to be analyzed is ACTTGAGAT, clicking the Reverse Sequence button converts it to TAGAGTTCA.

3.5 Translation

This function translates DNA sequences into amid acid.

Explanation of the Result Window

Bar for displaying a translation.

Map View

Sequence View

trace-scf   Sequence
            Translate

ACGACTTCCT CTTTCATTCA
ThrThrSerS erPheIleGl
ArgLeuPro LeuSerPheA
AspPheLe uPheHisSer

Result of translation
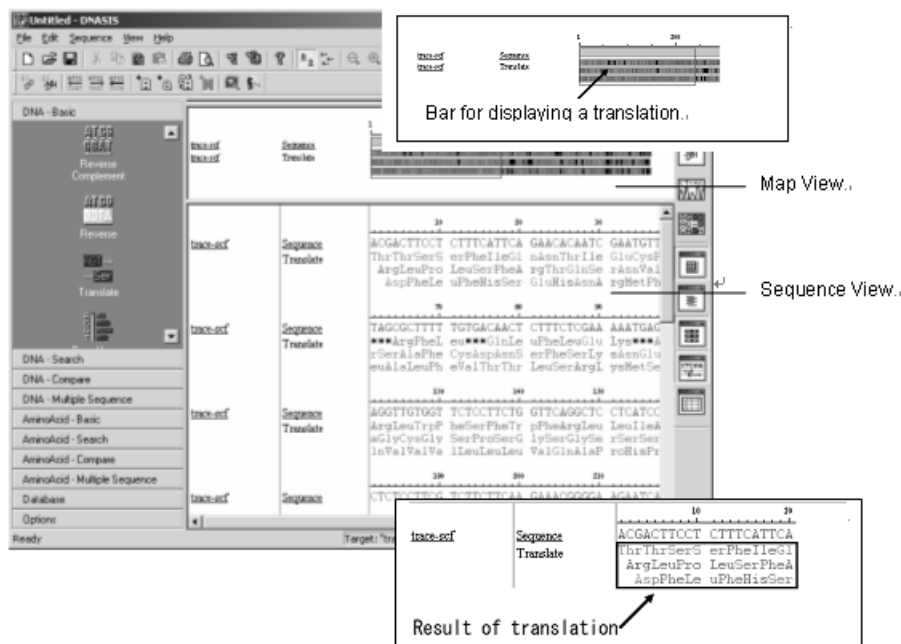
Map View

In the color display mode, this view displays a bar that shows the sequence in the colors of amino acid residues.   By

default, the property of amino acid is classified into four groups: acidic, basic, neutral (polar), and neutral

(nonpolar).   The color for each group is set as follows:

Red	Acidic	Asp	Glu					
Blue	Basic	Lys	Arg	His				
Yellowish green	Neutral (Polar)	Trp	Ser	Thr	Cys	Gln	Asn	Tyr
Green	Neutral (Nonpolar)	Ala	Val	Leu	Ile	Pro	Phe	Met	Gly
Black	Miscellaneous	Uncertain, stop codon						

Sequence View

This view displays the translated amino acid sequences in a three-row pattern for each frame.

*Refer to "5.8 Codon Table".

The sequences are translated according to the conversion rules in the Codon Table*.   The DNA sequences are

translated for each group of three characters, so that some bases may not be translated.

Order of translation→

GTC	GCC	AAG	CAC	AT
|	|	|	|	|
V	A	K	H	Not translated

This function translates a nucleic acid character string that differs from any of the combinations specified in the

Codon Table as follows:

1. Any replaceable characters are replaced.   The Codon Table is searched for all combinations to perform a translation.
2. If the combination does not match any entry in the Codon Table, that combination is translated into 'X'.
3. If more than one codon matches, the function checks whether all the amino acids translated as non-X are identical. If all are identical, they are translated into the same amino acid. If one of them is different, that is translated into X.

The method for replacing characters is as follows:

R -> G,A   M -> A,C   B -> G,T,C   V -> G,C,A   Y -> T,C   S -> G,C

D -> G,A,T   N -> A,C,G,T   K -> G,T   W -> A,T   H -> A,C,T

Example 1: Translating AAH

Because AAH is not found in the Codon Table, it is treated in the following way.   Of AAH, A is not a replaceable

character so that it is not replaced.   H can be replaced with A, C or T.   Accordingly, AAH can be replaced with any

of the following: AAH, AAA, AAC, and AAT.

AAH / AAA / AAC / AAT

Using the replaced character string, the function searches the Codon Table again to perform a translation.   The first

AAH does not exist.   The next AAA can be translated into the amino acid of K (No. 43 in the table).   Similarly,

AAC is translated into N (No. 42 in the table) while AAT is translated into N (No. 41 in the table).   Because all the

three results (K, N, N) are not the same amino acid, AAH is translated into X.

Example 2: Translating TCN

Because TCN is not found in the Codon Table, the character string becomes the target of translation.   Of TCN , T

and C are not replaceable characters so that they are not replaced.   N can be replaced with A, C, G or T.

Accordingly, TCN can be replaced with any of the following: TCN, TCA, TCC, TCG, and TCT.

TCN / TCA / TCC / TCG / TCT

Using the replaced character string, the function searches the Codon Table again to perform a translation.   The first

TCN does not exist.   The next TCA can be translated into the amino acid of S (No. 19 in the table).   Similarly,

TCC is translated into S (No. 18 in the table), TCG into S (No. 20 in the table), and TCT into S (No. 17 in the table).

Because all the four Ss are the same amino acid, TCN is translated into S.

Specifying a Frame to Display

1. In the Sequence View, right-click in the result of translation and select the Property menu.
2. A frame names are displayed in the Frame field in the Translate View Property window as shown in the figure.
   Place a checkmark in the check box of the frame you want to display.
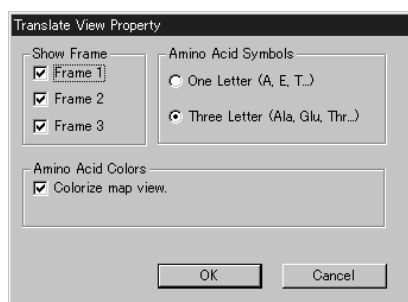
3. Click the OK button.

Changing to One-Character Notation

You can select a one-character or three-character notation to display the result of translation.

1. In the Sequence View, right-click in the result of translation and select the Property menu.

2. Select "One Letter" in the Amino Acid Symbols field in the Translate View Property window as shown in the figure.
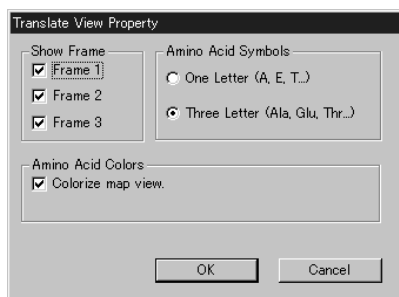
3. Click the OK button.

Changing Codon Table

*Refer to "5.8 Codon Table".

You can select a codon table for translation from the registered codon tables*.

1. Click the Translation icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and a Parameter dialog box will appear.

2. Select a codon table for translation in the Codon Table in the Translate window as shown in the figure.

3. Click the OK button.

Changing the Display Color of Amino Acids

You can change the display colors of amino acid sequences in the result of translation.   The initial setting provides four different colors.

1. Selects View and then Preference.... Alternatively, you can click the 🖳 button on the toolbar.

2. Select the Sequence tab in the Preferences window.

3. Select the Colorize Sequence check box then select Amino Acid.

4. Select amino acid in the combo box for selecting amino acid.
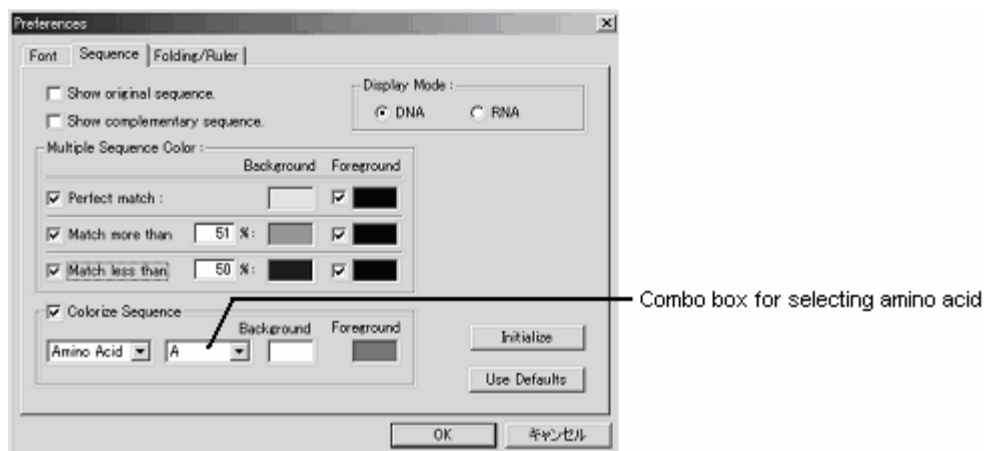
5. Double click Background or Foreground then set the color on the color palette.

6. Repeat Steps 4 and 5 as required.

7. Click the OK button.

Editing and Analyzing the Result of Translation

1. In the result of translation, drag the cursor to select the portion of the frame you want to edit or analyze, as shown in the figure.
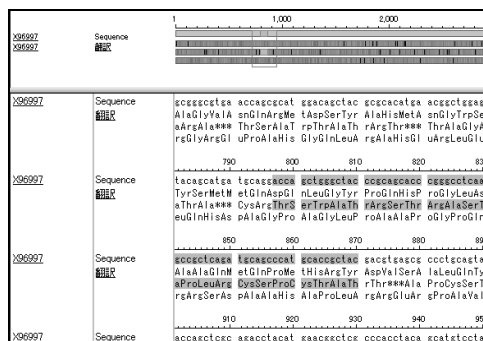
2. Click the ( ) button on the toolbar.

3. The Amino Acid window appears.   This window displays the amino acid sequence for the selected portion in the Sequence View, as shown in the figure.
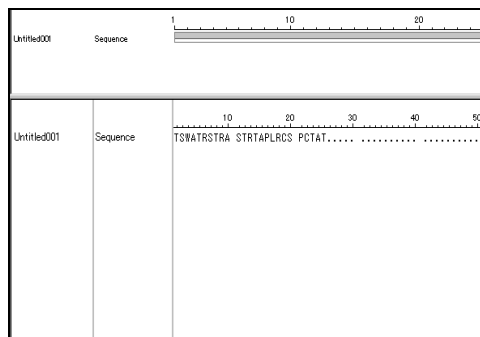
3.6 Base Content

This function analyzes and displays the ratio of bases that comprise DNA sequences.   The result of analysis is

displayed in another window.

Explanation of the Result Window
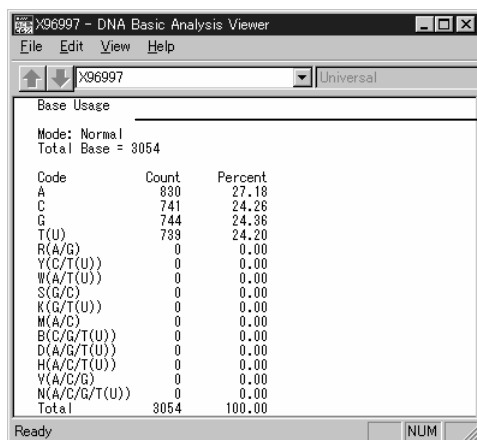
Analysis mode

File menu	Description
Export	Exports the data in the window into a text file.
Print	Prints the window.
Print Preview	Displays a printing image.
Print Setup	Provides various print settings.
Exit	Closes the result window.

Edit menu	Description
Copy	Copies the data in the window as a tabbed character string into the Clipboard.

View menu	Description
Toolbar	Toggles the toolbar to display/hide it.
Status bar	Toggles the status bar to display/hide it.

Help menu	Description
About DNABasicAnalysisViewer	Displays the version information of this analysis function in a dialog.
Contents	Displays online help.

Button		Description
🖫	Export button	Provides the same function as the Export menu.
🖨	Print button	Provides the same function as the Print menu.
🖹	Copy button	Provides the same function as the Copy menu.
❓	Help button	Displays online help.

3.7 Codon Usage

This function displays the type and number of codons in DNA sequences.   The result of analysis is displayed in another window.

Explanation of the Result Window

File menu	Description
Export	Exports the data in the window into a text file.
Print	Prints the window.
Print Preview	Displays a print image.
Print Setup	Provides various print settings.
Exit	Closes the result window.

Edit menu	Description
Copy	Copies the data in the window as a tabbed character string into the Clipboard.

View menu	Description
Toolbar	Toggles the toolbar to display/hide it.
Status bar	Toggles the status bar to display/hide it.

Help menu	Description
About DNABasicAnalysisViewer	Displays the version information of this analysis function in a dialog.
Contents	Displays online help.

Button		Description
💾	Export button	Provides the same function as the Export menu.
🖨	Print button	Provides the same function as the Print menu.
📋	Copy button	Provides the same function as the Copy menu.
❓	Help button	Displays online help.

Changing the Frame

1. Click the Translation icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and a Parameter dialog box will appear.

2. Select a frame in the Start Position field in the Codon Usage window as shown in the figure.

3. Click the OK button.

3.8 GC Content

This function calculates and analyzes the percentage of G or C that is included in every 10 bases of a DNA sequence.

The result of analysis is graphically shown in another window.

Explanation of the Result Window
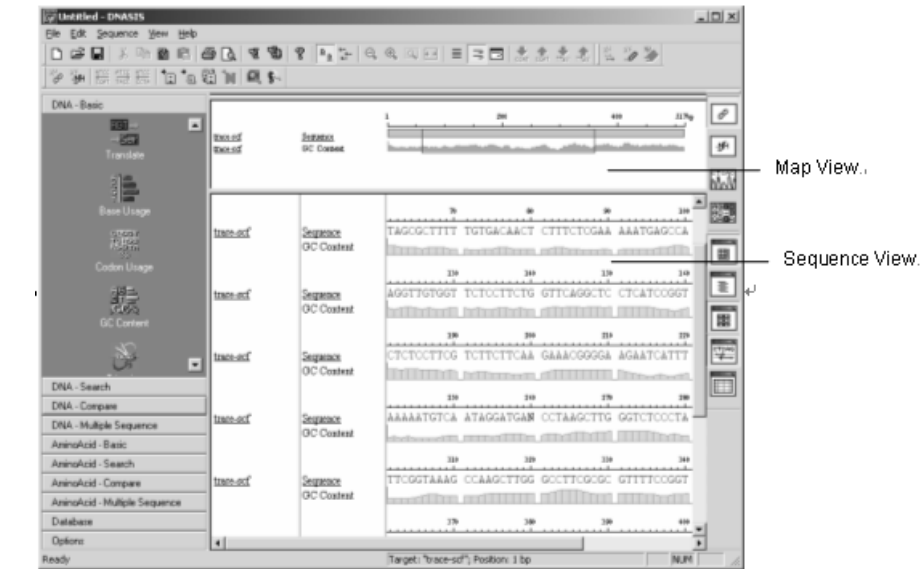
Map View

This view displays the graph of the entire sequence on the bar.

Sequence View

This views graphically displays the result below the sequence.

Customizing the Result Display

If you right-click in a graph, a menu is displayed. This menu is used to customize the form of graphs.
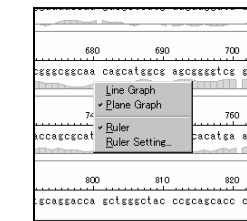
Menu	Function
Line Graph	Displays the GC content in a bar graph
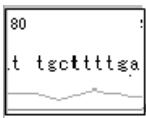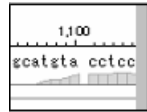
Plane Graph		Displays the GC content in a histogram (This is the initial setting).
Ruler		Displays the ruler. Use the ruler to obtain rough measurement. You can move the ruler vertically by gripping the ruler line using the mouse.
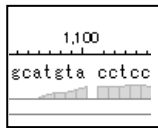
3.9 Vector and Low-Quality End Trimming

This function searches DNA sequences for low-quality end portions and vector sequences.   It also displays the region where the low-quality end portions and vector sequences have been trimmed.

Explanation of the Result Window
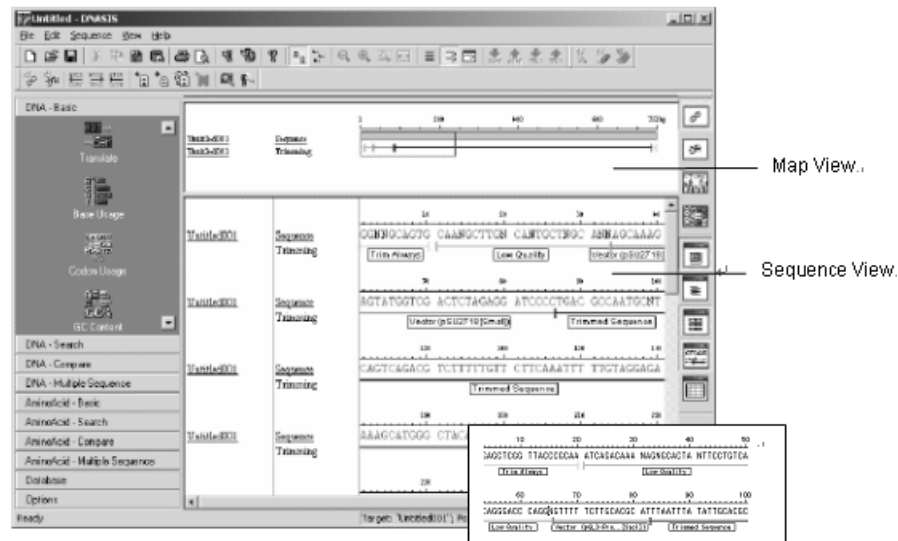
Sequence View

This view displays the result of trimming according to the preset conditions below the sequence.   It displays the following:

- Trim Always: Shows the region that has been unconditionally removed from the end irrespective of the quality or vector sequences.
- Low Quality: Shows the region that has been removed because of low quality.
- Vector xxxx: Shows the region that has been removed as the vector sequences.
- Trimmed Sequence: Shows the region from which the vector sequences and the low quality region have been removed.

Trimming Only Vectors

In the initial setting, both the vectors and the end are trimmed.   You can change the initial setting to trim only the vectors.

1. Click the Vector & Low Quality Trim End icon and an Analysis dialog box will appear. Click the Parameter button and a Parameter dialog box will appear.
2. Click the Trim End checkbox to uncheck it in the Parameterset Editor window as shown in the figure.
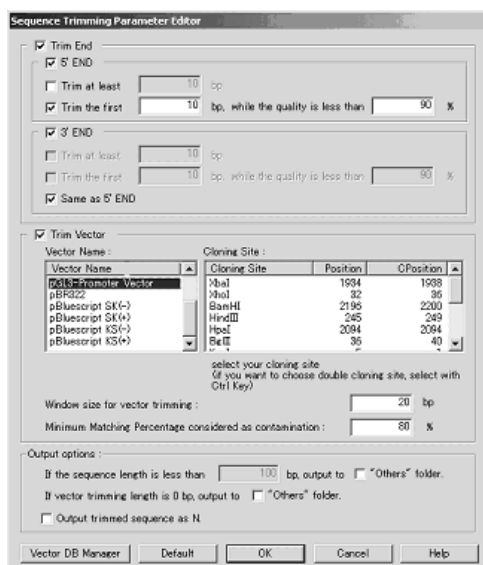
3. Select the vector you want to trim from the Vector name list in Trim Vector.

4. Select a cloning site.   You can select up to two cloning sites by pressing the Ctrl key.

5. Click the OK button.

Registering New Vectors

In addition to vectors registered in DNASIS in advance, you can register new vectors for trimming in the vector

database.

1. In the Analysis Button View, click the Database tab and then Vector Database to open the Vector Database
   Manager window.

2. Click New... at the bottom of the Vector Database Manager window to display the New Vector window as shown
   in the figure.
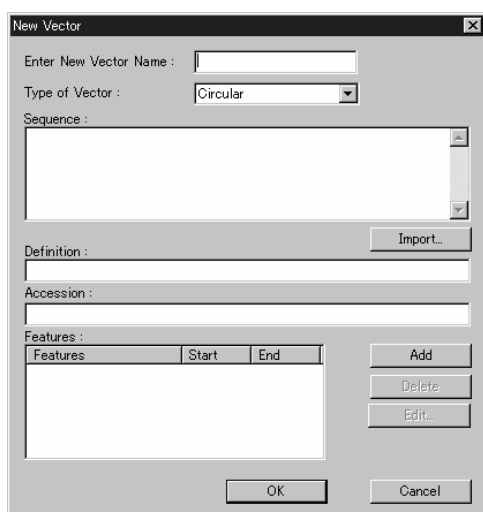
3. When you create a new vector, enter an appropriate vector name in the Enter New Vector Name field.   If you use
   an existing vector, click Import... first to open a dialog box and specify the file you want to import.

4. You must fill in the Vector Name, Type of Vector, and Sequence fields.   Fill in the other fields as required.
   Complete all the settings and then click the OK button.

5. The new vector has been added to the Vector Database Manager.   The cloning site is automatically set.   Confirm
   the contents and click the OK button.

Trimming Low-Quality End

1. Click the Vector & Low Quality Trim End icon and an Analysis dialog box will appear. Click the Parameter button and a Parameter dialog box will appear.

2. Set the trimming conditions on the 5' end. Place a checkmark in three checkboxes: Trim end, 5' END, and Trim the first.
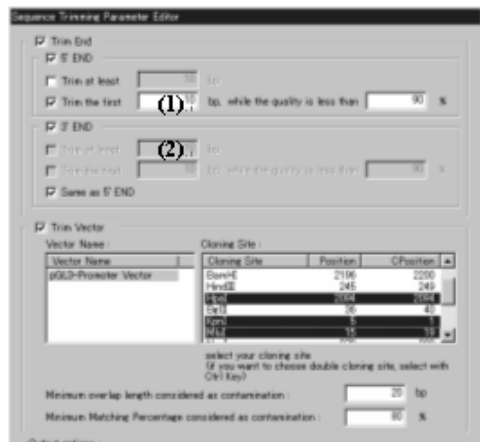
Enter a numeric value in the (1) field (10 in the example). If the range to calculate quality is 10, it means that the quality value is calculated every 10bp count.

Enter a numeric value in the (2) field. This value shows a criterion to determine whether or not the quality is low.

3. Click the OK button.

Trimming Unconditional End

1. Click the Vector & Low Quality Trim End icon and an Analysis dialog box will appear. Click the Parameter button and a Parameter dialog box will appear.
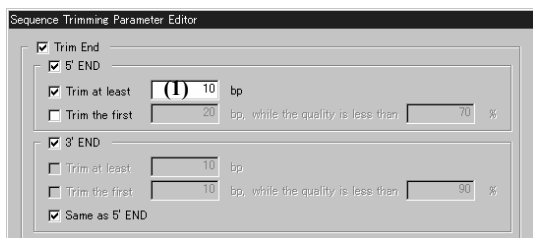
2. Enter the bp count of the end to be trimmed in the (1) field in the Parameterset Editor window as shown in the figure.

3. Click the OK button to close the dialog box.

Analyzing the Trimmed Sequence

There are two ways to analyze the trimmed sequence.

Selecting the Trimmed Sequence

1. In the result of analysis for trimming, click the bar indicated with Trimmed Sequence. A trimmed sequence is now selected as shown in the figure.
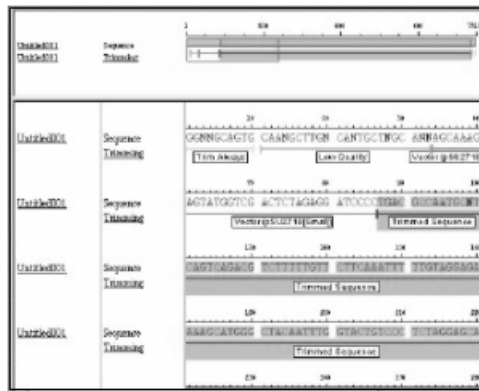
2. Click the ![button] button on the toolbar.

3. The trimmed sequence is now added as a new DNA sequence.

Replacing the Trimmed Part with N

1. In the analysis result for trimming, click all the bars indicated with Trim Always, Low Quality, and Vector xxx to

   select them.   If you want to select more than one item, click the second and subsequent items by pressing the Ctrl

   key.

2. Click the ![button] button on the toolbar.

3. The trimmed portion is now replaced with N.

3.10 ORF

This function searches DNA sequences for open reading frames (ORF) and displays the result.

Explanation of the Result Window

Sequence View

This view displays the result of searching for ORFs together with the sequences.

The | symbol on the bar indicates a start codon and the > symbol indicates a stop codon.   If you click an ORF

between the start and stop codons, the ORF is selected and highlighted by the predefined color.

Map View

This view displays the result of searching for ORFs under the sequences.

The □ symbol on the bar indicates a start codon and the > symbol indicates a stop codon.   If you click an ORF

between the start and stop codons, both the ORF and the sequence in the region are selected.

Changing the Codon Table

1. Click the ORF icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and a Parameter dialog box will appear.
2. Select the codon table you want to change in the Codon Table in the Parameterset Editor window as shown in the figure.   To verify or edit the codon table, click the ... button.
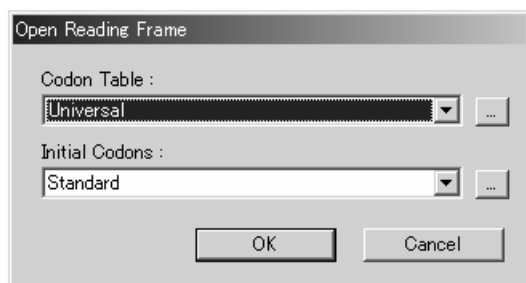
3. Click the OK button.

Changing the Start Codon

1. Click the ORF icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and a Parameter dialog box will appear.

2. Select the start codon you want to change in the Initial Codons field in the Parameterset Editor window as shown in the figure.   To verify or edit the start codon, click the ... button.
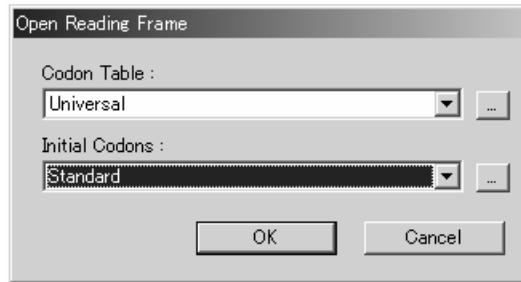
3. Click the OK button.

Listing the Result of a Search for ORFs

For open reading frame result in sequence view, select the sequence name and analysis name then click the Result List Dialog button.   All the ORFs are listed:

You can copy and save the ORF list.   The data that has been copied or saved can be used by other applications such as MS-Excel.

Click ▸▸ to display a list of Start Codons and Stop Codons not in the reading frames.

Selecting an ORF to Display

1. Follow the procedure in the previous operation to display the ORF List window.

2. Click the checkboxes of the ORFs you do not want to display in the list to uncheck them.   Immediately after analysis, the start and stop codons outside the reading frame are not displayed in the Map View.   However, you can display them by placing a checkmark on this list.

3. Click the OK button.

Narrowing Down the ORFs to Display

1. In the Sequence View, double-click in the result of searching for ORFs.   Alternatively, you can right-click an ORF in the Sequence View to select the Show Setting menu.

2. Select the frames you want to display from the Frame field.   Those frames with the checkmarks placed in the checkbox are displayed.

> Note: If you select the All Frame field, all ORFs are displayed in a single frame.

3. Select the length of an ORF you want to display from the ORF field.
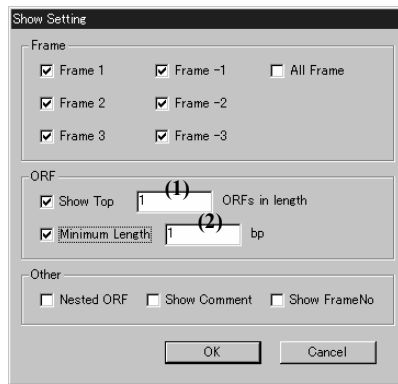
(1): Specify the number of ORFs you want to display, starting with the longest one.

(2): Specifies the length of the shortest ORF you want to display.

If you place checkmarks in both 1 and 2, only the ORFs that meet both conditions are displayed.

4. Sets the following in the Other field.

Nested ORF: Forcibly draws the starting point if nested ORF is displayed.

Show Comments: Displays the comments for the ORF in the Sequence View.

Show FrameNo: Displays the frame numbers in the Sequence View.

5. Click the OK button.

Adding a Selected ORF Sequence to the Editor

1. In the Sequence View or Map View, click an ORF to select it.

2. Click the ![icon] button on the toolbar with the ORF selected.

3. DNA sequence for the selected ORF is now added in the Sequence View, so that you can continue to analyze DNA sequence of the ORF.

Adding a Comment to a Selected ORF

1. Double-click an ORF to which you want to add a comment.   The ORF Shape Setting window appears as shown in the figure.

2. Enter a comment you want to display in the Comment field.

3. Click the OK button.   A comment appears under the ORF as shown in the figure.
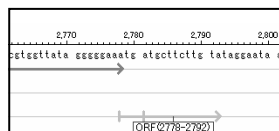
Creating Amino Acid Translated Sequence for an ORF

1. Click the ORF button to search for ORFs.

2. Click the Translate button to start translation.

3. Click the desired ORF.   Then, the result of translation is selected.

4. Click the Amino Acid Transfer Button on the toolbar .

 The window switches to the Amino Acid mode.   A new amino acid sequence representing the desired ORF is

 created in amino acid view of this project.

3.11 Primer Design

This function designs PCR primers.

Explanation of the Result Window

Map View.

Sequence View.

Product   Primer Sequence   Result of Primer Design

Sequence View

If you click the primer or product, the sequence is selected.   If you double-click the primer, its detail is displayed in a dialog box.

Displaying the Primer List

For the primer design result in sequence view, select the sequence name and analysis name then click the Result List Dialog button.

Selecting the Primer That Amplifies a Selected Range

If you know the region you want to amplify in a sequence in advance, you can design the primer to always include this region in the PCR product.

1. In the Sequence View, select a portion you want to amplify.

2. In the Analysis Button View, click Primer Design.

Range to be amplified

Primer that amplifies a selected range

Selecting a Primer to Display

1. In the Sequence View, right-click in the result of primer design and select the Show Result List menu.   For the primer design result in sequence view, select the sequence name and analysis name then click the Result List Dialog button.   A window that indicates the list of results of primer design appears.
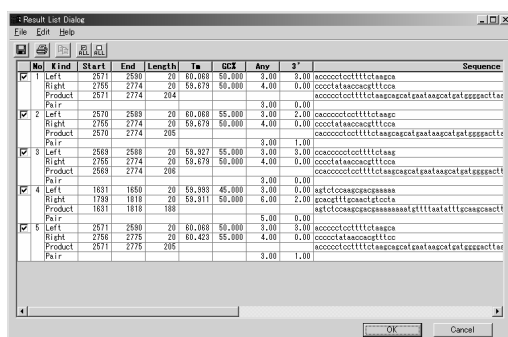
2. Click any of the check boxes on the left to uncheck a primer you do not want to display.

3. Click the OK button.

Changing the Tm Value for a Primer to be Designed

1. Click the Primer Design icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and a Parameter dialog box will appear.

2. Select the Primer Picking Conditions tab in the Parameterset Editor window as shown in the figure.
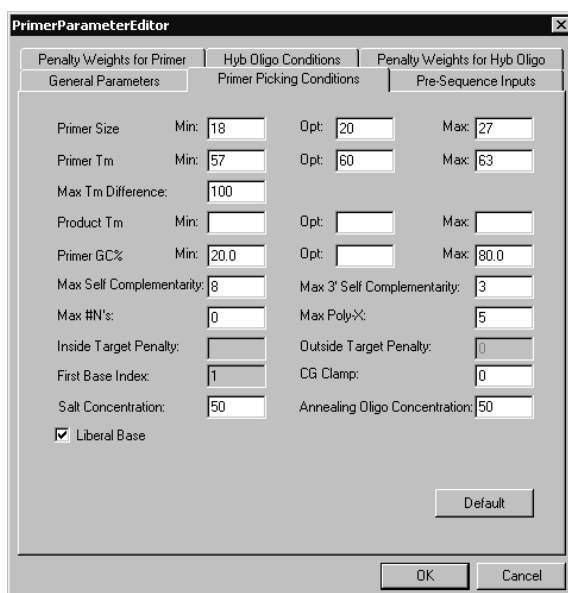
3. Set the following values in the Primer Tm field.

Min: The minimum Tm value for the primer to be designed.

Note: The primers whose Tm values are smaller than this value cannot be designed.

Opt: The primers whose Tm values are as close to this value as possible are designed.

Max: The maximum Tm value for the primer to be designed.

Note: The primers whose Tm values are larger than this value cannot be designed.

4. Click the OK button.

Changing the Length for a Primer to be Designed

1. Click the Primer Design icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and a Parameter dialog box will appear.

2. Select the Primer Picking Conditions tab in the Parameterset Editor window as shown in the figure.

3. Set the following values in the Primer Size field.

Min: The minimum length for the primer to be designed.

Note: The primers whose length is shorter than this value cannot be designed.

Opt: The primers whose length is as close to this value as possible are designed.

Max: The maximum length for the primer to be designed.

Note: The primers whose length is longer than this value cannot be designed.

4. Click the OK button.

Exporting the Result to Excel

1. For the primer design result in sequence view, select the sequence name and analysis name then click the Result List Dialog button.
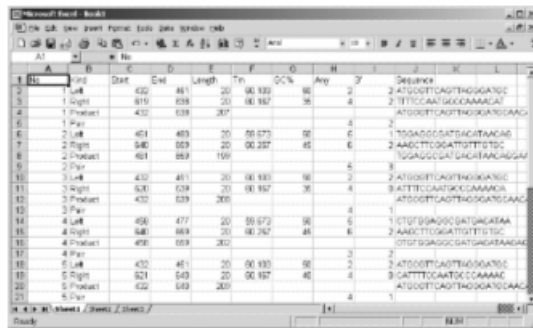
2. To copy the whole list into Excel, simply save as text file and open the file in Excel.

3. Alternatively, Select Copy All or Copy Selected Cells from Edit in the menu.

Copy All: Copies all the information being displayed.

Copy Selected: Copies only the columns that have been selected.

4. Paste the copy to an MS-Excel sheet.

3.12 Oligo-Probe Design

This function designs oligo-probe for DNA sequences.

Explanation of the Result Window
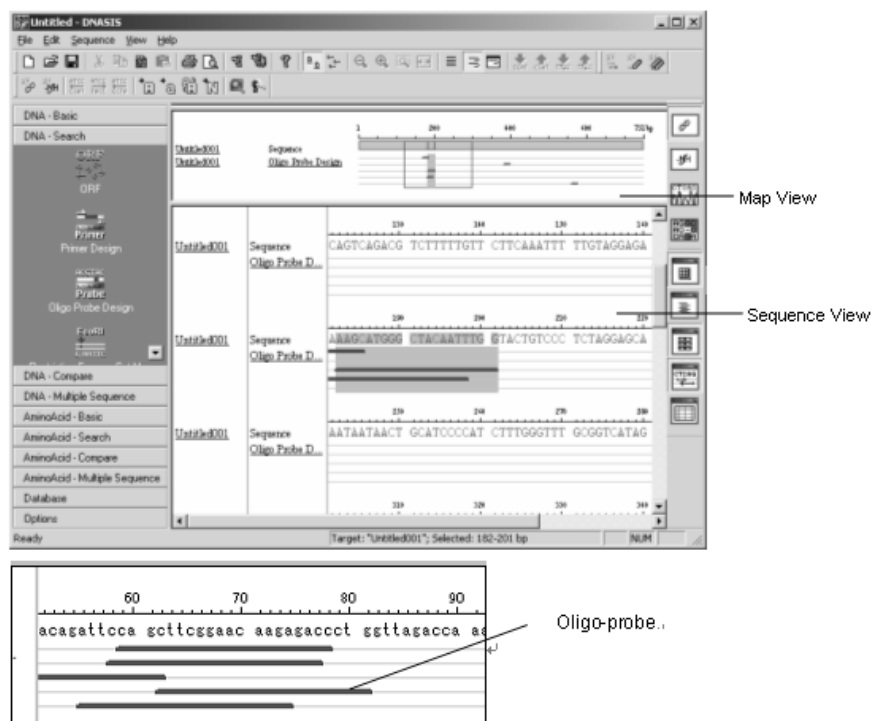
The designed probe is displayed on the bar.   You can change the number of probes that you want to display. If you

click the probe portion, the corresponding sequence is selected.

Displaying a List of Probes

For the oligo probe design result in sequence view select the sequence name and analysis name then click the Result

List Dialog button.

Exporting the Result to Excel

1. For the primer design result in sequence view, select the sequence name and analysis name then click the Result
   List Dialog button.
2. To copy the whole list into Excel, simply save as text file and open the file in Excel.

3. Alternatively, Select Copy All or Copy Selected Cells from Edit in the menu.

    Copy All: Copies all the information being displayed.

    Copy Selected: Copies only the columns that have been selected.

Paste the copy to an MS-Excel sheet.

Designing a Probe in a Specified Region

1. Click the Oligo Probe Design icon from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.

2. On the Pre-Sequence Inputs page, enter an appropriate value in the Included Region field.   Use the following format: "<start bp>, length".   (In the example, the design is based on a length (up to 149bp) from 50bp to 100bp.)

3. Click the OK button.

3.13 Restriction Site Search

This function searches DNA sequences for recognition and cutting sites of restriction enzymes and displays the result of search.

Explanation of the Result Window
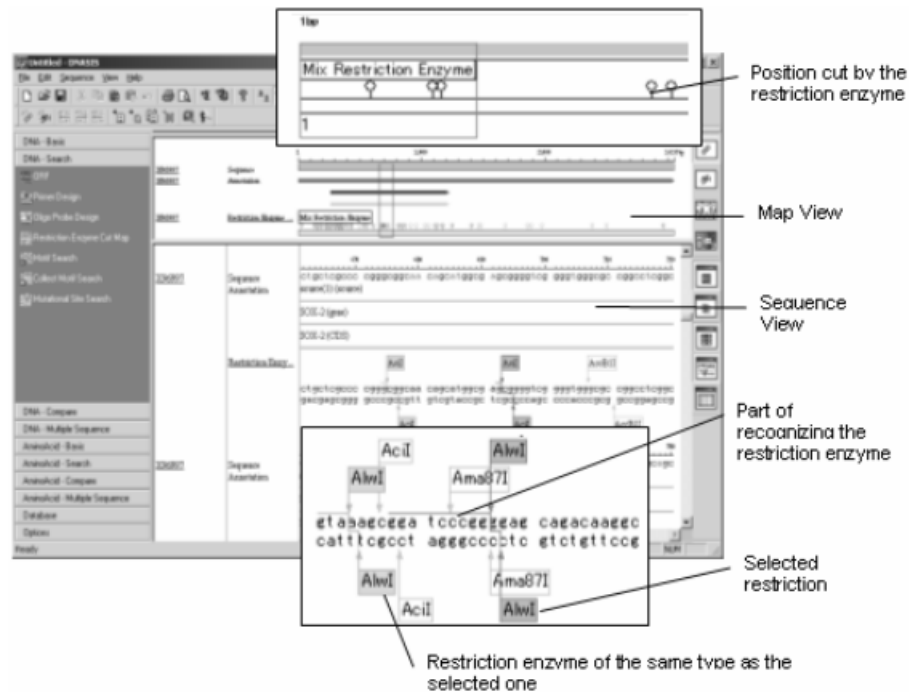
Map View

The pin shows the position where the restriction enzyme cuts.   If you move the cursor to the pin and click it, the display color* changes and the pin is selected.   If there is more than one position to cut by the same restriction enzyme, all of them are highlighted.

*You can change the display color. Refer to "1.5 Preferences Dialog Box".

Sequence View

Together with the sequence, the following are displayed: the name of restriction enzyme, the part of recognition, and the position to cut.   If you click the mouse, the part to cut the restriction enzyme is selected.   If there is more than one position to cut by the same restriction enzyme, all of them are highlighted.

The part displayed in a red frame in the Map View is displayed in the Sequence View.

Selecting a Restriction Enzyme to be Searched for

The restriction enzymes are registered in the restriction enzyme database.   The position to cut by the selected restriction enzyme is searched for from the database.

1. Click the Restriction Enzyme Site Search icon from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.   The RestrictionSiteParamEditor window appears.

2. When searching for the length of recognition sequence or a cutoff, select the Select from a Category item and check the target Recognition Length and Cut Kind. If necessary, designate the upper and lower limits for cut frequency. When selecting a target restriction enzyme from the list, select the Select from a List item, and click ... Restriction enzymes registered in the Restriction Enzyme List will appear. Also, when selecting restriction enzymes from the list, it is possible to designate the upper and lower limits of cut frequency.
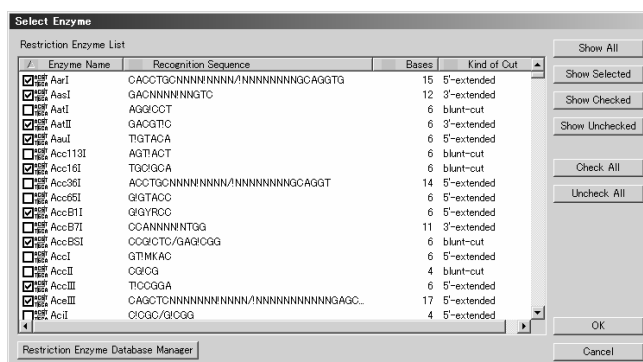
3. The ones with a check in the box to the left of the Enzyme Name are the selected restriction enzymes. Select the restriction enzyme to search, and click OK.

4. Click OK on the RestrictionSiteParamEditor window.

Registering a New Restriction Enzyme

1. Click the Restriction Enzyme Site Search icon from analysis button view. The Analysis dialog box will appear then click the Parameter button to display the Restriction Site ParamEditor window.

2. Select the Select from a List item then click ... to display the Select Enzyme dialog.

3. If you want to register an existing restriction enzyme, select Import to select a file you want to register.
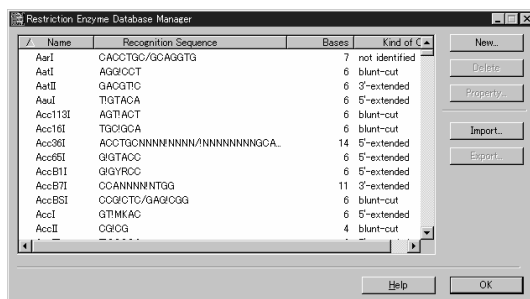
When you want to create a new restriction enzyme, select New....   The New Enzyme window appears as shown in the figure.   Enter the required items and click the OK button.
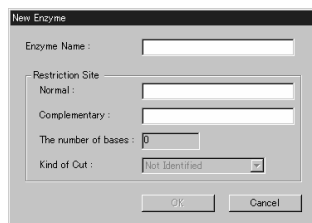
Selecting a Restriction Enzyme to Display

1. For the enzyme restriction result in sequence view, select the sequence name and analysis name then click the Result List Dialog button.
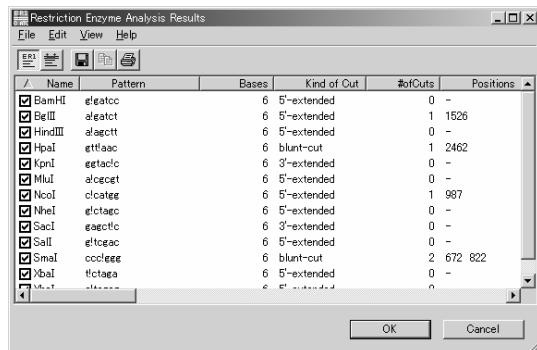
2. After the search completes, the restriction enzymes that cut the sequence are listed.   Place a checkmark for the restriction enzyme you want to display and click the OK button.

Selecting a Sequence that Contains a Cut Piece

If you click the region between two sites in the Map View area, the fragment is displayed in the predefined color and the sequence of that fragment is also selected.
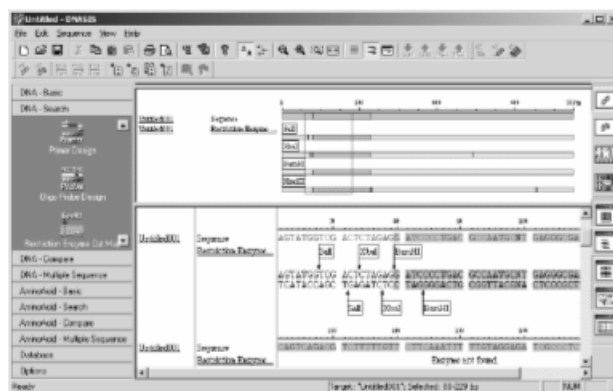
Looking for a Restriction Enzyme That Cuts Out a Specified Range

1. In the sequence, select the range you want to cut out with the shortest overhanging length.
2. In the Sequence View, right-click and select Search Optimum Enzyme... from the menu.   The Search Optimum Enzyme Options window appears.
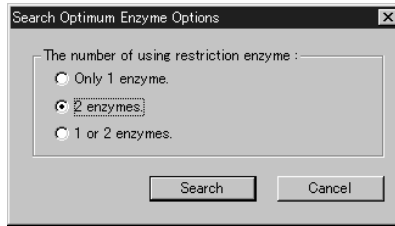
Only 1 enzyme.	Searches for an optimum restriction enzyme that cuts out the selected range using one type of restriction enzyme.
2 enzymes.	Searches for an optimum restriction enzyme that cuts out the selected range using two types of restriction enzyme.
1 or 2 enzymes.	Searches for an optimum restriction enzyme to cut out the shortest length using one type of restriction enzyme. Alternatively, uses two types of restriction enzyme in case that is more effective.

3. Specify any of the choices and click the Search button to search for a restriction enzyme.

4. The found position to cut is displayed in the predefined color.

Display Restriction Enzyme Fragment List

Displays a list of fragments obtained by cutting the target sequence at the point of restriction enzyme sites.

To display the restriction enzyme fragment list, select the sequence name and analysis name then click the Result

List Dialog button in Map View or Sequence View, or click 🗮 in the Restriction Enzyme Fragment List dialog.

3.14 Motif Search

This function searches a DNA sequence for one or more motif(s).   There are two ways of specifying the motifs to be searched for: one using a database and the other by keying in any sequence pattern.

Explanation of the Result Window
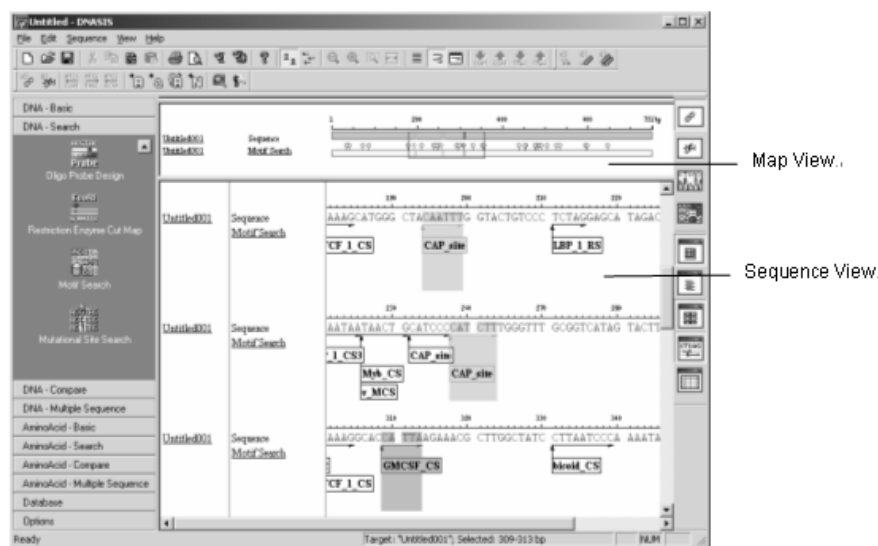
Map View

The pin shows the found motif.   If you move the cursor to the pin and click it, the display color changes and the pin goes into the selected status.

Sequence View

Together with the sequence, the following are displayed: the name of the motif and the part of recognition.   If you click the motif name, it is displayed in the predefined color and the sequence that contains the motif also goes into the selected status.

Searching for Motifs listed in a Database

1. Click the Motif Search icon in DNA Search from analysis button view and an Analysis dialog box will appear.
   Click the Parameter button and the Nucleic Acid Motif Search Parameter Set Editor window appears.
2. Place a checkmark for the Use Motif Database and select the database displayed in the list.
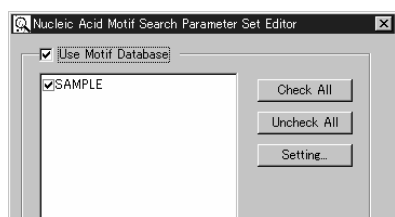
3. Select the OK button.

Searching for a specific sequence Motif

1. Click the Motif Search icon from analysis button view and an Analysis dialog box will appear. Click the
   Parameter button and the Nucleic Acid Motif Search Parameter Set Editor window appears.

2. Place a checkmark for the Use Input Pattern.   Enter or paste a motif you want to search for.

3. Click the OK button.

The Analysis Result View shows a motif that has the name "Input Pattern".

Displaying a List of Search Results

For the frame in sequence view select the sequence name and analysis name then click the Result List Dialog button.

Use Copy All or Copy Selected from Edit in the menu to copy all or selected cells to the clipboard as tab delimited text.

Use Save All as or Save Selected as from File in the menu to store all or selected cells in a file as tab delimited text.

Adding a Motif Database

You can add a new database to the motif database.   For details, refer to "Adding a Motif Database" in "5.5 Amino Acid Motif Database".

Browsing the Detail of the Found Motif

In the Sequence View or Map View, double-click the motif to display its details.

Nucleic Acid Motif Annotation

Name NF_E1_CS1

Pattern MYWATCWY

Range 15 bp – 22 bp Strand Normal

Annotation

```
*********
NF_E1_CS1
*********
This motif recognize sequence pattern "MYQATCWY".
This motif has no annotation.
```

OK

3.15 Mutation Site Search

This function searches for restriction enzyme recognition positions, such that on a per-frame basis, a one-base replacement will not affect the result of translation but prevent a cut by that restriction enzyme.   The resulting DNA sequence with information about such potential "mutation sites" is displayed in a new window.

Explanation of the Result Window

		Codon Table Name	Universal					
Frame No.	Enzyme NAME	Find Position	Mutation Seq.	MUT Pos.	Recognize Seq.	CUT Pos.	Translation	
3	EcoHI	188 – 192	gcggg	188	!ccsgg	187	ccc–>ccg : P	
2	EcoHI	208 – 212	gccgg	208	!ccsgg	207	cgc–>cgg : R	
3	EcoHI	230 – 234	gccgg	230	!ccsgg	229	ccc–>ccg : P	
2	EcoHI	238 – 242	gccgg	238	!ccsgg	237	ggc–>ggg : G	
1	EcoHI	294 – 298	gccgg	294	!ccsgg	293	ccc–>ccg : P	
2	EcoHI	313 – 317	tcggg	313	!ccsgg	312	ctc–>ctt : L	
2	EcoHI	358 – 362	gcggg	358	!ccsgg	357	cgc–>cgg : R	
3	EcoHI	428 – 432	gcggg	428	!ccsgg	427	ccc–>ccg : P	
3	EcoHI	599 – 603	accgg	599	!ccsgg	598	atc–>ata : I	
1	EcoHI	651 – 655	gcggg	651	!ccsgg	650	ctc–>ctg : L	
3	EcoHI	698 – 702	tcggg	698	!ccsgg	697	ggc–>ggt : G	
2	EcoHI	709 – 713	gccgg	709	!ccsgg	708	gcc–>gcg : A	
2	EcoHI	721 – 725	gcggg	721	!ccsgg	720	gcc–>gcg : A	
3	EcoHI	1079 – 1083	tccgg	1079	!ccsgg	1078	ccc–>cct : P	
1	EcoHI	1119 – 1123	gccgg	1119	!ccsgg	1118	gtc–>gtg : V	
1	EcoHI	1227 – 1231	gccgg	1227	!ccsgg	1226	ggc–>ggg : G	

Codon Table name	Codon table name used
Frame No.	Frame number
Enzyme NAME	Name of the restriction enzyme
Find Position	Position of the restriction enzyme searched for
Mutation Seq.	Sequence of the mutation site
MUT Pos.	Position of the mutation
Recognize Seq.	Recognition sequence for the restriction enzyme
CUT Pos.	Position to cut
Translation	Change in translation by the mutation site

Selecting a Codon Table

1. Click the Mutation Site Search icon from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.

2. Select a codon table you want to change from the Codon Table field in the Mutation Site Parameter Editor window as shown in the figure.
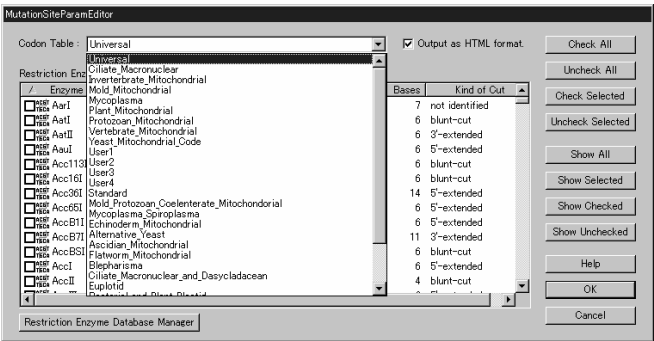
3. Click the OK button.

Selecting a Restriction Enzyme

1. Click the Mutation Site Search icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and the Mutation Site Parameter Editor dialog box will appear.
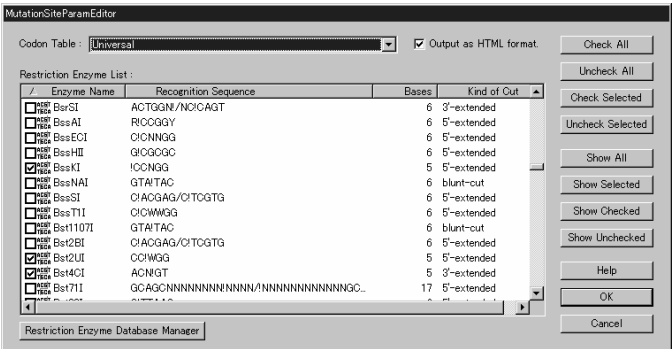
2. The Restriction Enzyme List lists the restriction enzymes registered.   The Enzyme Name with a checkmark specified in the check box on the left is the selected restriction enzyme.   Select a restriction enzyme that is searched for in the mutation site.

3. Click the OK button.

3.16 Hairpin Loop Search

Searches and displays the results of hairpin loop position for a DNA or RNA sequence selected from the sequence editor.

Explanation of the Result Window

Map View

Sequence View

Hairpinloop(40)

Stem Area   Stem Area
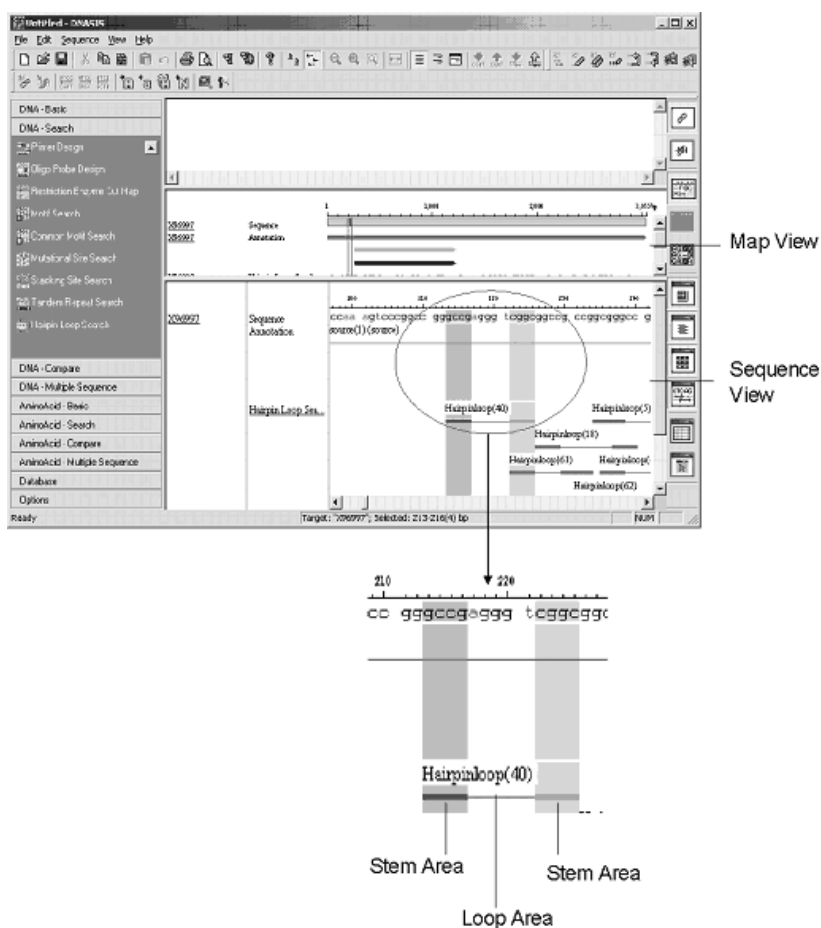
Loop Area

Map View

Displays total hairpin loop areas.   Click a stem area to color it as selected and another stem area as highlighted. Click a loop area to color the total hairpin loop area as selected.

Sequence View

Displays the sequences together with the stem/loop areas.   Click a stem area to color it as selected and another stem area as highlighted.   Click a loop area to color the total hairpin loop area as selected.   Sequences are selected or highlighted linked areas.

Displaying a List of Search Results

For analysis result in sequence view, select the sequence name and hairpin loop name then click the Result List Dialog button.
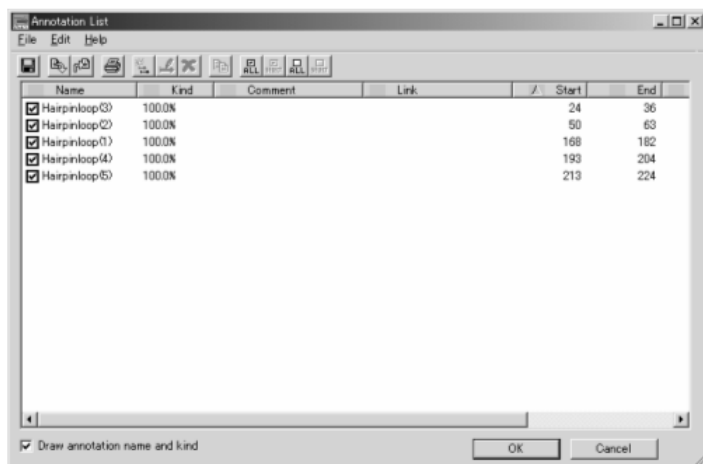
In the dialog, the list of hairpin loop areas can be copied, saved, printed and so on.   For detail, refer to "Annotation

List Dialog" in "4.32 Annotation".

Setting Parameters

1. Click the Hairpin Loop Search icon from analysis button view.   An Analysis dialog box will appear then click
   the Parameter button and the Hairpin Loop Parameterset Editor will appear.
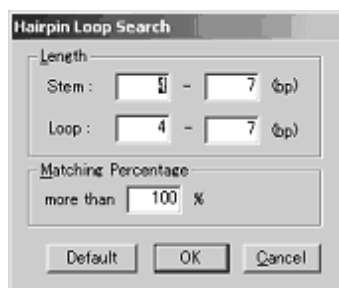
2. Set the Length and Matching Percentage columns.

3. Click the OK button.

3.17 Stacking Site Search

Searches and displays the results of stacking site position for a DNA or RNA sequence selected from the sequence editor.

Explanation of the Result Window
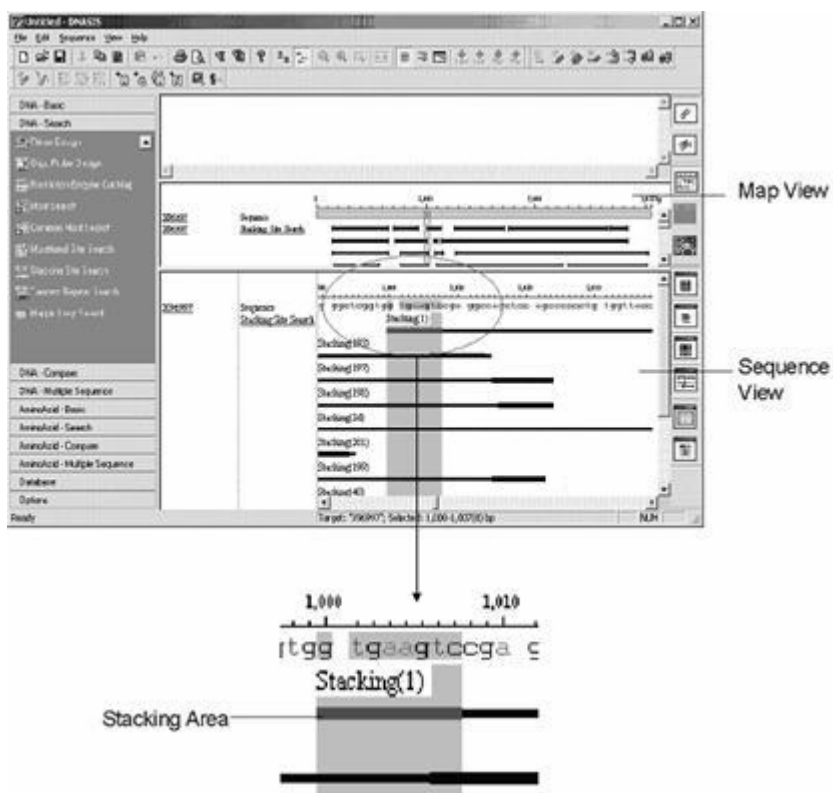
Map View

Displays stacking areas. Click a stacking area to color it as selected and another area as highlighted.

Sequence View
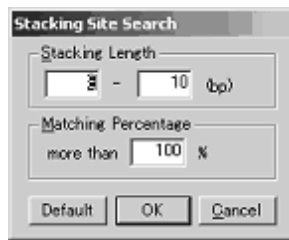
Displays the sequences together with the stacking areas.   Click a stacking area to color it as selected and another area as highlighted.   Sequences are selected or highlighted linked areas.

Displaying a List of Search Results

Refer to "Displaying a List of Search Results" in "3.16 Hairpin Loop Search".

Setting Parameters

1. Click the Stacking Site Search icon from analysis button view.   An Analysis dialog box will appear then click the Parameter button and the Stacking Parameterset Editor will appear.

2. Set the Stacking Length and Matching Percentage columns.

3. Click the OK button.

3.18 Tandem Repeat Search

Searches and displays the results of tandem repeat position for a DNA or RNA sequence selected from the sequence editor.

Explanation of the Result Window
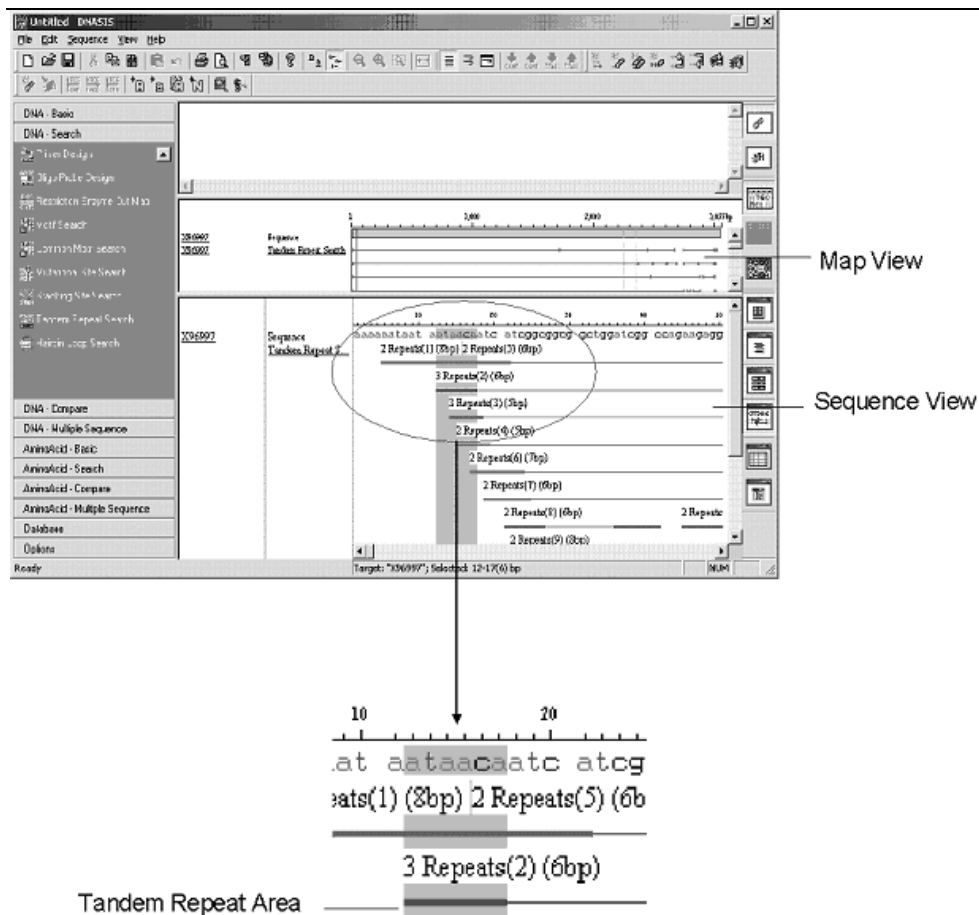
Map View

Displays tandem repeat areas.   Click a tandem repeat area to color it as selected.

Sequence View

Displays the sequences together with the tandem repeat areas.   Click a tandem repeat area to color it as selected.

Displaying a List of Search Results

Refer to "Displaying a List of Search Results" in "3.16 Hairpin Loop Search".

Setting Parameters

1. Click the Tandem Repeat Search icon from analysis button view.   An Analysis dialog box will appear then click the Parameter button and the Tandem Parameterset Editor will appear.

2. Set the Repeat Length and Repeat Count columns.

3. Click the OK button.

3.19 BLAST Search

This function uses the BLAST algorithm to perform a homology search between a DNA sequence and the specified BLAST database.   The result of search is displayed in another window.

Types of BLAST Search

There are four types of BLAST search for DNA sequences.

Button name	Program name	Description
BLAST search	blastn	Homology search between DNA sequences and a DNA sequence database.
BLAST search (Protein DB)	blastx	When you enter a DNA sequence, performs an all-frame translation and then a homology search between amino acid sequences and an amino acid sequence database.
BLAST search (Translation DB)	tblastx	When you enter a DNA sequence, performs an all-frame translation and then a homology search between amino acid sequences and the DNA sequence database that has been translated for all frames.
One-to-One BLAST Search	blastn	Performs a one-to-one BLAST search between two DNA sequences.

Explanation of the Result Window

The following explains how to operate the menu in the result window.
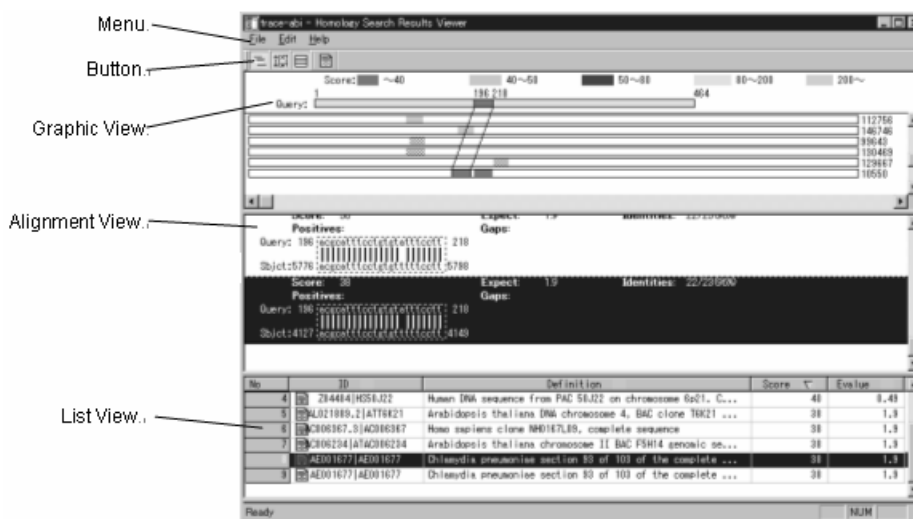
File menu	Description
Export Alignment	Exports content in alignment view to a file.
Export List	Exports content in list view to a file.
Print Setup	Display Print Property Dialog.
Print Preview	Display print preview for selected data in the view.
Print	Print the data selected in the view.
Print All	Print all data in the view.
Exit	Closes View.

Edit menu	Description
Copy	Copies the selected portion of PairwiseView.

Help menu	Description

| Contents | Displays online help. |
| Homology Search Results Viewer | Displays the version information. |

Button	Description
≡	Toggles the Graphic View to display/hide it.
ACGT / ACAT	Toggles the Alignment View to display/hide it.
☰	Toggles the List View to display/hide it.
▤	Display the sequence selected in the view as GenBank Report in the external viewer. (To display the sequence in GenBank Report Viewer, the corresponding "Space" is needed.)
🖶	Print the data selected in the view.
🔍	Display print preview for selected data in the view.

Graphic View

 - If you click a matched region of the Subject, the region is selected and the matched part of the Query is displayed

    in a color influenced by the Score.

 - If you click the white part in the window, the selection is canceled.

 - By clicking the mouse while pressing the Shift key, you can select more than one matching region.

Explanation of Window Images

Query Part (Top of the window)

 - The numeric value indicated at the top of the bar in the initial status refers to the Query range.

 - In the selected status, the range of a match is indicated by the numeric value at the top of the bar.

Subject Part (Bottom of the window)

 - One sequence corresponds to one bar.

 - The numeric value at the rightmost of the bar indicates the sequence length.

 - A match is displayed in the color corresponding to Score.

    (A shaded part represents a complement sequence.)

 - If a single sequence contains more than one match, the same bar displays these matches.   The highest match in

    terms of Score is aligned with the Query to serve as the reference position.   Each of the other matches is

    displayed in a relative position from the reference position.

 - The gray bar indicates the correct length.

 - The white bar indicates the length longer than the window by fixing the width of non-matching parts.

Alignment View

 - Displays all alignments.

 - Double-click the icon on the left of the sequence header to obtain the sequence's GenBank/Report and add it to the

    DNASIS Main window.

 - The background of sequences in the selected status is displayed in the Windows-based color.

- Shows the Match sequences between Query sequence and the Subject sequence. (BLAST searches for protein, translation and amino acid databases)

Item name (Parameter name)	Description
Type	Shows the original database where the subject sequence has been registered (gb: GenBank, emb: EMBL, dbj: DDBJ, etc.).
ID	Shows the ID of the entry in the original database where the subject sequence has been registered.
Length	Shows the length of the subject sequence.
Score	Shows the score of a match. A match with a higher score value is higher in similarity.
Expect	Shows the expected value of a match. A match with a lower score value is higher in similarity.
Identities	Shows the percentage of the matching bases (or amino acids) within the entire length of a match.
Positives	Shows the number of groups in which the score has a positive value within the entire length of a match when the query sequence and the subject sequence are compared for each amino acid.
Gaps	Shows the total number of gaps inserted into the query sequence and the subject sequence. This cell remains blank when there is no gap.

List View

By default, sorting is carried out in descending order in terms of Score.   To switch sort items, click the title part (each time you click, the order alternately changes between descending and ascending.   Once the sort item is switched through a title click, the current sort item becomes the second sort item.

Explanation of Window Images

Item name (Parameter name)	Description
No	Line number
ID	Shows the ID of the entry in the original database where the subject sequence has been registered.
Definition	Provides a brief description of sequences.
Score	Shows the score of a matching part. Any matching part with a higher score value has higher similarity.
Evalue	Shows the expectation of a matching part. Any matching parts with a lower score value has higher similarity.

Selecting a Database to Be Searched (other than one-to-one BLAST Search)

1. Click the BLAST Search icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and BLAST Parameters will appear.
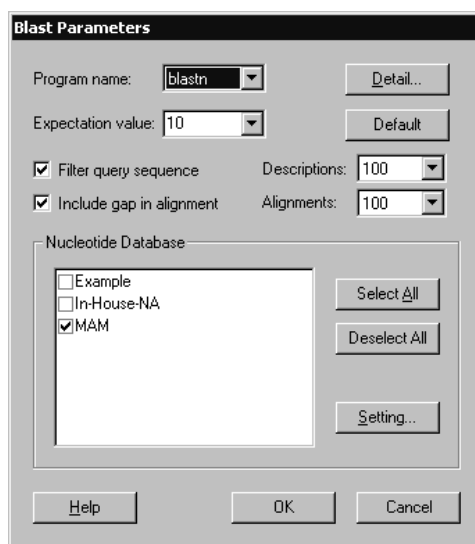
2. The Nucleotide Database filed displays a list of databases.   Place a checkmark for the database to be searched.

3. Click the OK button.

Obtaining an Entry to the Result of Search

If the object entry for the result of search belongs to the GenBank database, it is possible to obtain the entire

GenBank Flat file of the entry.   Since this function links to the NCBI Web site via the Internet, the Internet

environment and the proxy server must be set.

Procedure:

Select an entry you want to obtain and click the [icon] button on the toolbar.   You can select more than one entry by

clicking the mouse while pressing the Shift key.

3.20 Internet BLAST Search

This function provides homology search using the BLAST search service from the NCBI Web site.   For the

analysis, establish an Internet environment*.

Types of BLAST Search

There are three types of BLAST search for DNA sequences.

Button name	Program name	Description
BLAST search	blastn	Homology search between DNA sequences and a DNA sequence database.
BLAST search (Protein DB)	blastx	When you enter a DNA sequence, performs an all-frame translation and then a homology search between amino acid sequences and an amino acid sequence database.
BLAST search (Translation DB)	tblastx	When you enter a DNA sequence, performs an all-frame translation and then a homology search between amino acid sequences and the DNA sequence database that has been translated for all frames.

Explanation of the Result Window

Refer to "3.19 BLAST Search".

Selecting a Database to Be Searched

1. Click the BLAST Search icon from analysis button view and an Analysis dialog box will appear. Click the
   Parameter button and the Internet BLAST Search Parameterset window will appear.
2. Click Setting... to display the NCBI Advanced BLAST Search window.

3. From the Database Selection combo box, select a database you want to search.
4. Click the OK button.

Selecting the Species

1. Click the BLAST Search icon from analysis button view and an Analysis dialog box will appear. Click the
   Parameter button and the Internet BLAST Search Parameterset window will appear.
2. Click Setting... to display the NCBI Advanced BLAST Search window.

3. From the Database Selection combo box, select a database you want to search.

4. Click the OK button.

3.21 Smith-Waterman Search

This function provides high-precision homology search using the Smith-Waterman algorithm between the input sequence and the target database. [The optional GENE BRIGHT III board allows this rigorous similarity search to occur at a high speed on a personal computer.]

Types of Smith-Waterman Search

There are two types of Smith-Waterman search for DNA sequences.

Button name	Description
Smith-Waterman search	Performs a Smith-Waterman search between a DNA sequence and a DNA sequence database.
One-to-One Smith-Waterman Search	Performs a Smith-Waterman search between two different DNA sequences.

Explanation of the Result Window

Refer to "3.19 BLAST Search".

Selecting a Database to Be Searched (Smith-Waterman search only)

1. Click the Smith-Waterman Search icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and a GENE BRIGHT III Parameterset Editor will appear.
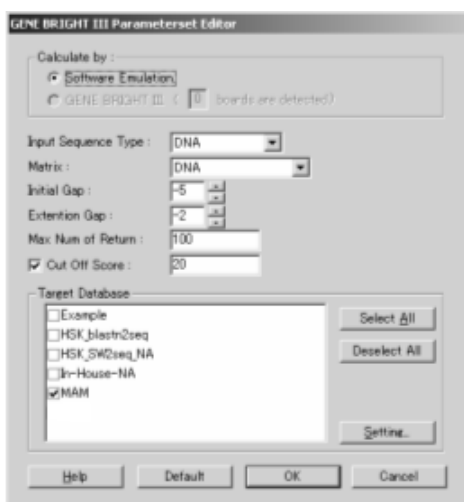

2. In the Target Database field, place a checkmark for the database to be searched.

3. Click the OK button.

3.22 Multiple Alignment

This function calculates and displays a multiple alignment of all sequences displayed in the window (ie. Sequences in a project) using the Clustal W algorithm.

Explanation of the Result Window

Map View.
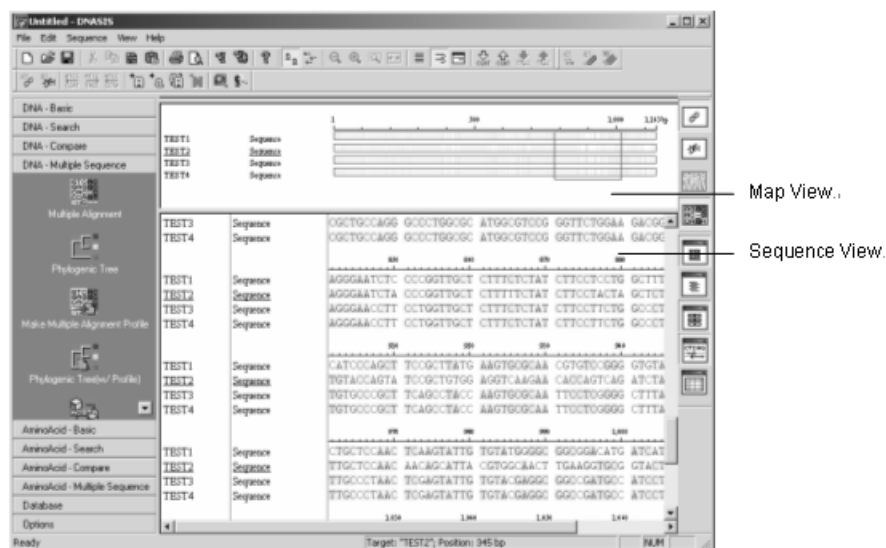
Sequence View.

Map View

This view displays matching conditions in the entire alignment.   If you move the cursor, the Map View also moves accordingly.

Sequence View

This view displays the alignment according to the perfect match, partial match, and non-match.   By default, yellow is for 100% matches, green for matches of 51% or more, and light blue for matches of 50% or less.   The portion in a red frame in the Map View is displayed in the Sequence View.

Example of Calculation Time

The following is an example of calculation time without using bootstrap on a Pentium III 550MHz machine.

Average sequence length	Number of sequences				
	100	200	300	400	500
100	0:00:17	0:01:40	0:05:45	0:15:50	0:32:25
200	0:01:21	0:05:28	0:14:29	0:31:20	0:58:27
400	0:05:06	0:20:24	0:48:37	1:35:55	2:31:56
1000	0:30:21	2:02:05	4:24:51	7:55:38	12:34:42
1500	1:08:00	4:32:36	10:22:05	17:34:58	30:23:36

Analyzing a Selected Range with Sequence Analysis Tools

In the alignment display mode, you cannot directly perform other types of analysis. To carry out sequence analysis on regions of sequence based on multiple alignment view perform the steps below.

1. In Alignment view, select a region you want to analyze as shown in the figure.

2. On the Switch Pane Toolbar click the Sequence mode so that sequence analysis and annotation entry tools become active. Now you can analyze of the region of the alignment you highlighted in step 1.

Meaning of the Background Color and How to Change It

The result of a multiple alignment is color-coded according to the degree of conservation of the individual bases across all aligned sequences.

You can change the matching rate and color combination.

1. Click in the View-Preference menu or click the 🔲 button on the toolbar to open the Preference dialog box.

2. Click the Sequence tab.
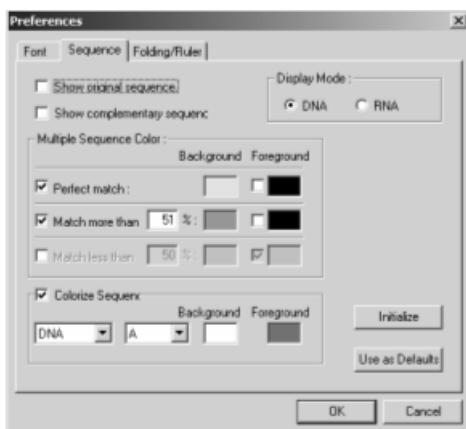
Perfect match	Colored when the bases of all sequences match.
Match more than	Colored when the matching rate is higher than a preset value.
Match less than	Colored when the matching rate is lower than a preset value.

3. Set the parameters under Multiple Sequence Color.

4. Click the OK button.

Editing an Alignment Sequence

You can introduce as well as edit gaps within sequences when they are part of an alignment.   Use the "-" mark to introduce a gap.

*Refer to "Changing the Order of Sequence Display" in "3.22 Multiple Alignment".

Changing the Order of Sequences

After the initial alignment is displayed, you can move sequences up or down in the order to improve the display according to your needs.

Choosing Sequences in a Project that are to be Aligned

You can use a data list if you want to align only those specified sequences among the sequences displayed in the window.

1. On the Switch pane tool bar, click the data list 🔲 button to display the list of sequences in the project.
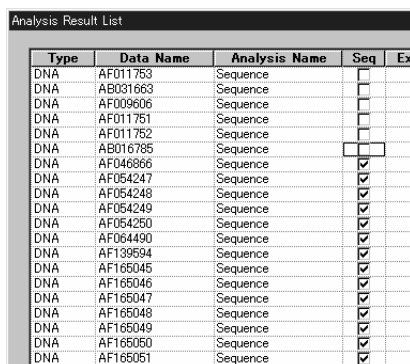
2. Uncheck the Seq field of the sequences you do not want to add.
3. Click the OK button.
4. Click the Multiple Alignment button to start analysis.   (Perform an analysis after canceling the alignment display mode.)

Alignment after Masking Regions within Sequences

1. Select region within one or more sequences and press the 🔲 button on the toolbar.   The selected region is shown by an N and masked as shown in the figure and is not considered in calculating the alignment.

2. Click the Multiple Alignment button to start the analysis.
3. The analysis will be performed without including the masked part as shown in the figure.

Creating a Consensus Sequence

This function uses the result of multiple alignment, to create and display a consensus sequence based on the most conserved base at each position in the alignment. In the drop down menu of the result window under "Sequence", select "Make Consensus".   The consensus sequence is added to the Sequence View.

```
                              10        20        30        40        50
                         |    .    |    .    |    .    |    .    |    .    |
M0001      Sequence      TGCCCGCCCC TAAATGGGG- CGACACTCCG CCATGAATCA CTCCCCTGTG A
M0002      Sequence      -GCCAGCCCC CTGATGGGGG CGACACTCCA CCATAGATCA CTCCCCTGTG A
M0003      Sequence      -GCCAGCCCC CTGATGGGGG CGACACTCCA CCATGAATCA CTCCCCTGTG A
M0004      Sequence      -GCCAGCCCC CTGATGGGGG CGACACTCCA CCATGAATCA CTCCCCTGTG A
Untitled001 Consensus    TGCCAGCCCC CTGATGGGGG CGACACTCCA CCATGAATCA CTCCCCTGTG A

                              70        80        90        100       110
                         |    .    |    .    |    .    |    .    |    .    |
M0001      Sequence      GTCTTCACGC AGAAAGCCGTC TAGCCATGGC GTTAGTATGA GTGTCGTACA G
M0002      Sequence      GTCTTCACGC AGAAAGCCGTC TAGCCATGGC GTTAGTATGA GTGTCGTGCA G
M0003      Sequence      GTCTTCACGC AGAAAGCCGTC TAGCCATGGC GTTAGTATGA GTGTCGTGCA G
M0004      Sequence      GTCTTCACGC AGAAAGCCGTC TAGCCATGGC GTTAGTATGA GTGTCGTGCA G
Untitled001 Consensus    GTCTTCACGC AGAAAGCGTC TAGCCATGGC GTTAGTATGA GTGTCGTGCA G

                              130       140       150       160       170
                         |    .    |    .    |    .    |    .    |    .    |
M0001      Sequence      CCCCCCCTCC CGGGAGAGCC ATAGTGGTCT GCGGAACCGG TGAGTACACC G
M0002      Sequence      CCCCCCCTCC CGGGAGAGCC ATAGTGGTCT GCGGAACCGG TGAGTACACC G
M0003      Sequence      CCCCCCCTCC CGGGAGAGCC ATAGTGGTCT GCGGAACCGG TGAGTACACC G
M0004      Sequence      CCCCCCCTCC CGGGAGAGCC ATAGTGGTCT GCGGAACCGG TGAGTACACC G
Untitled001 Consensus    CCCCCCCTCC CGGGAGAGCC ATAGTGGTCT GCGGAACCGG TGAGTACACC G
```

3.23 Phylogenic Tree- DNA

This function calculates a phylogenic tree by using all sequences that are currently displayed in the window.   The result of calculation is displayed in another window.

Explanation of the Result Window
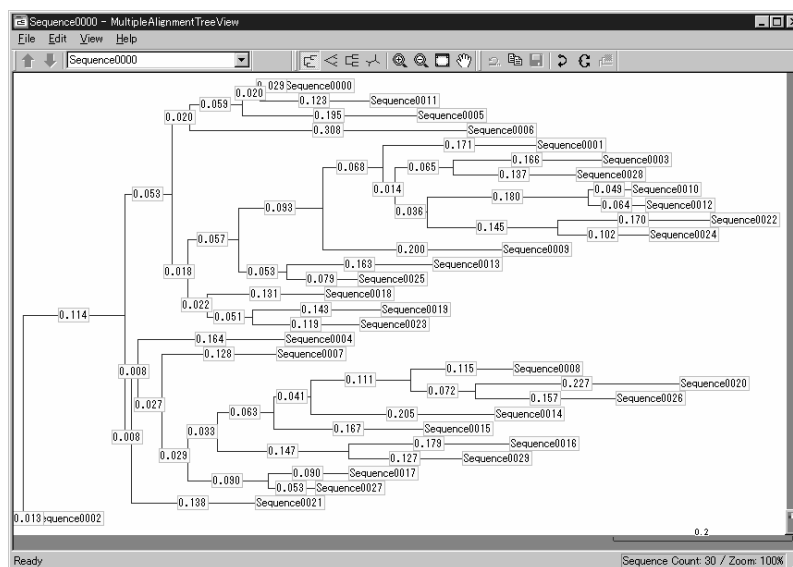
This result window uses the Phylogram format where the sequence name is displayed on the right at the end of each horizontal line and the evolutionary distance is displayed at each branch point.   The length of each horizontal line is proportional to the evolutionary distance.

File menu	Explanation
Export...	Saves input data as an external file that is given a name.
Export Tree...	Names and stores the phylogenic tree data in DND (dendrogram) format.
Save	Saves the currently displayed data by overwriting the original data. Note: At present, this function is not available.
Save as...	Saves the currently displayed data by using a different filename. Note: At present, this function is not available.
Print...	Performs printing.
Print Preview	Confirms the image of printing. If you click Close, you can exit from the Print Preview mode and return to the original display mode.
Print Setup...	Sets the size of printing paper.
Exit	Closes the window.

Edit menu	Explanation
Undo	Cancels the previous edit operation.
Copy	Copies the image of a phylogenic tree into the Clipboard.
OutGroup	Starts the edit command "Set an Out-Group".
Swap Branch	Starts the edit command "Exchange Branches".
Emphasis	Starts the edit command "Set Shading". Note: At present, this function is not available.

View menu	Explanation
Toolbar	Toggles the toolbar to display/hide it.
Status Bar	Toggles the status bar to display/hide it.
Phylogram	Changes the phylogenic tree display format to "Phylogram".
Slanted cladogram	Changes the phylogenic tree display format to "Slanted cladogram".
Rectangular cladogram	Changes the phylogenic tree display format to "Rectangular cladogram".
Unrooted	Changes the phylogenic tree display format to "Unrooted".
Zoom In	Increase the display size. Enlarge up to 1000%.
Zoom Out	Decrease the display size. Shrink down to 50%.
Whole indication	Display the phylogram according to the window size.
Preferences...	Displays a window for setting parameters.

Help menu	Explanation
Contents	Displays help for the Multiple Alignment Tree Viewer.
About MultipleAlignmentTree View	Displays the version information about the Multiple Alignment Tree Viewer.

Edit Tree Toolbar

Icon	Explanation
	The same as the Undo selection in the Edit menu.
	The same as the Copy selection in the Edit menu.
	The same as the Save selection in the Edit menu.
	The same as the OutGroup selection in the Edit menu.
	The same as the Swap Branch selection in the Edit menu.
	The same as the Emphasis selection in the Edit menu.

Tree View Toolbar

Icon	Explanation
	The same as the Phylogram selection in the View menu.
	The same as the Slanted cladogram selection in the View menu.
	The same as the Rectangular cladogram selection in the View menu.
	The same as the Unrooted selection in the View menu.
	The same as the Zoom In selection in the View menu.
	The same as the Zoom Out selection in the View menu.
	The same as the Whole indication selection in the View menu.
	Move the phylogram by dragging it.

Changing the Type of a Phylogenic Tree

You can select a phylogenic tree from four types: Phylogram, Slanted cladogram, Rectangular cladogram, and

Unrooted.   From the Tree View toolbar, select any type you want to display.
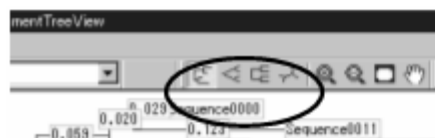
Phylogram

Slanted cladogram

Rectangular cladogram
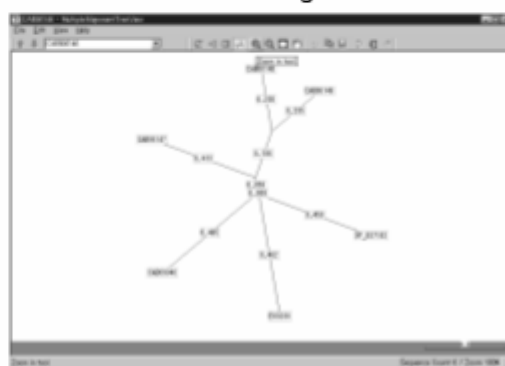
Unrooted.

Changing the Font

1. Select View-Preferences... to display the Parameter Set Editor.
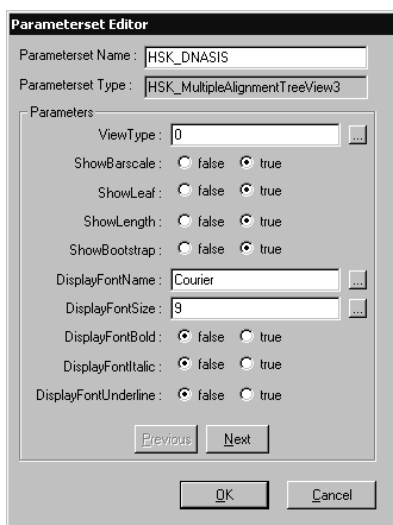
2. To change the font, use the Display Font Name field.   To change the size, use the Display Font Size field.

3. At the end of the setting operation, click the OK button to display a phylogenic tree in a new setting.

Displaying a Magnified Phylogenic Tree

1. Click the 🔍 icon on the toolbar to make the mouse cursor look like a magnifying glass.

2. Click or drag any section you want to magnify.   The specified section can be expanded.

To reduce it, click the ![icon] button and perform a similar operation.   You can return the displayed item to its original

size by clicking the ![icon] button.

Setting an Out-Group

You can set a selected branch as an out-group.

1. Click the ![icon] icon on the toolbar to change the mouse cursor to the + mark.

2. Move the cursor onto a branch you want to set to an out-group and click it.   The specified branch has now been
   set in the out-group.

Exchanging Branches

You can exchange branches.

1. Click the ![icon] icon on the toolbar to change the mouse cursor to the + mark.

2. Move the cursor to a branch you want to exchange with another within a tree and click it.   The specified branch
   is replaced and displayed.

Evaluating the Branching Reliability (Bootstrap Tree)

This function evaluates the reliability of a tree form using the bootstrap method.

1. Click the Phylogenetic Tree icon from analysis button view and an Analysis dialog box will appear. Click the
   Parameter button and a Multiple Alignment Parameter Editor will appear.
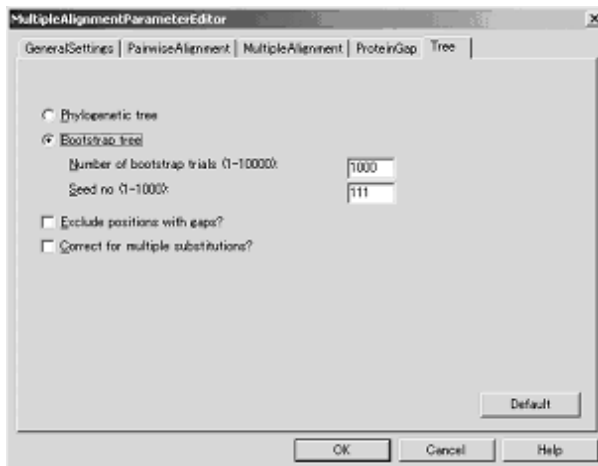
2. Click the Tree tab.

3. Select the Bootstrap tree.

     Number of bootstrap trials: The number of random trials

     Seed No: This is the random number seed for the random number generator.

Set these parameters.

4. Click the OK button.

5. Click the Phylogenic Tree button to start analysis.

3.24 Create a Phylogenic Tree for Manually Edited Alignments

After editing the contents in Alignment View, a new phylogenic tree is created from the results.

Procedure

1. Click Multiple Alignment & Phylogenic Tree in the Analysis Button View.

2. The Phylogenic Tree appears.

Result Window Description
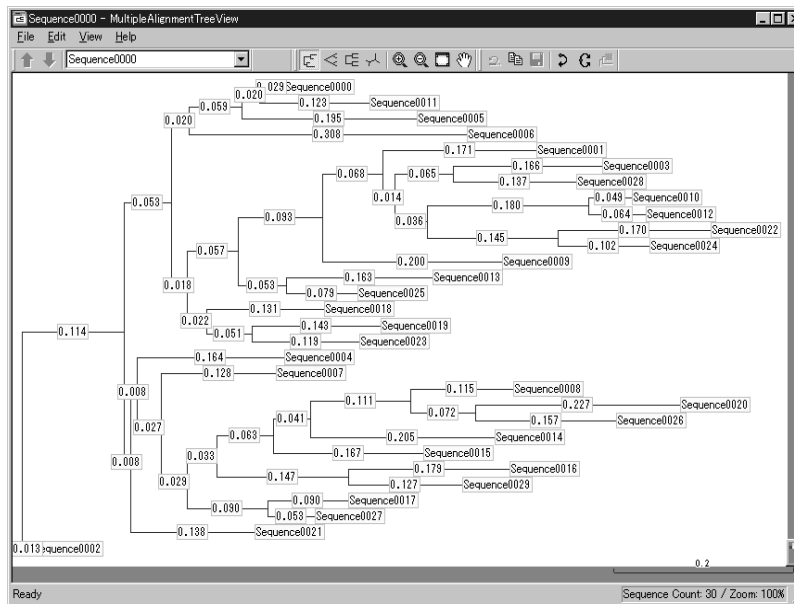
Refer to "3.23 Phylogenic Tree- DNA" for details.

3.25 Creating Multiple Alignment Profiles

This function creates a profile for a multiple alignment.   The multiple alignment between input sequences is calculated in advance and saved as a profile.   This allows high-speed alignment calculation between an unknown sequence and the profile.   The Clustal W method, developed by J. Thompson and T. Gibson, is used as an engine for alignment calculation.

What is a profile?

A multiple alignment profile is pre-calculated data for the alignments between multiple input sequences that is saved for later use.

Why do I want to use a profile?

Calculating multiple alignments is computationally intensive and can require a long time to complete.   DNASIS requires only ten minutes to calculate multiple alignments for 40 data items, but it may require two days for 200 data items.   This applies when the average length for the input DNA sequences is about 1.5Kbp.   Longer sequences, such as a gene or a complete genome, require a longer time.

If you have many known sequences and want to calculate alignment between an unknown sequence and the known ones, you can save the time required to calculate alignment with the unknown sequence by creating a profile (alignment of the known sequences) first.

Calculating a profile requires the same time as a typical multiple alignment.   However, once a profile is created, DNASIS can calculate alignment with the unknown sequence much faster (in about 10 seconds for the above example).

Disadvantages of using a profile

Using a profile provides fast calculation.   However, it results in degraded alignment precision.   The same set of sequences may produce different results when you use a profile and do not use a profile.   You should consider the above mentioned caveats when using a profile.

Procedure for Creating a Profile

As for any other analysis, click the Analysis menu and the tab DNA –Multiple Sequence. Note that this option is available only when in Sequence mode.

1. Enter or import sequences you want to align in order to create a profile into the Main window.
2. Click the Make Multiple Alignment Profile button and an Analysis dialog box opens.   Then click the Parameter button.
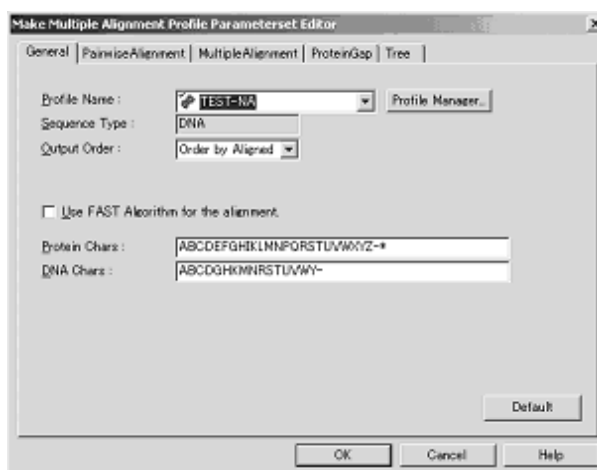
3. In the Profile Name field, select a profile you want to create, and click the OK button or to create a new profile, select Profile Manager.*

4. Click the Create Multiple Alignment Profile button.

DNASIS uses all sequences displayed in the Sequence View to perform multiple alignment, and then writes the result into the profile.

Note: Note if you had chosen an existing profile name, the profile is overwritten.   Locking the profile prevents an unexpected overwrite.   Use the Profile Manager* for locking the profile.

*Refer to "5.7 Multiple Alignment Profile".

Using a Created Profile on Another PC

You can export a newly created profile and save it to a file.   You can also import such an exported profile to use it on another PC.

Export Procedure

1. Click the Create Multiple Alignment Profile analysis menu and an Analysis dialog box.   Then click the Parameter button.
2. Click the Profile Manager... button to display the Multiple Alignment Profile Manager window.
3. From the display, select a profile you want to export.
4. Click the Export... button.   The following window appears.
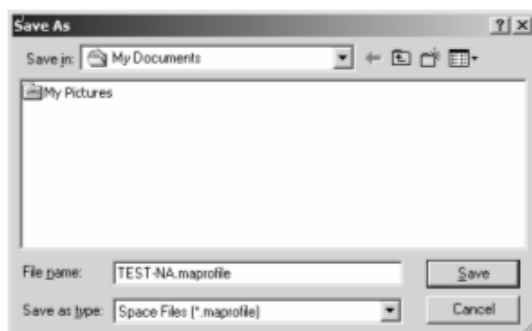
5. Specify the names of a folder and a file you want to save and click the Save button.

Import Procedure

1. Click the Create Multiple Alignment Profile analysis menu and an Analysis dialog box.   Then click the Parameter button.
2. Click the Profile Manager... button to display the Multiple Profile Manager.

3. Click the Import... button.   The following window appears.
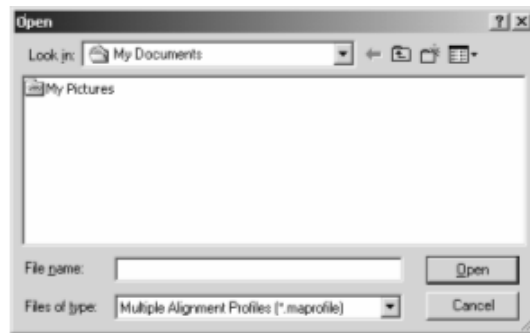
4. Specify the name of a file you want to import and click the Open button.   The imported profile is displayed in the list of the Multiple Alignment Profile Manager.

3.26 Using Phylogenic Trees - Profiles (DNA)

This function creates a phylogenic tree by adding a single sequence to a multiple alignment profile that has been produced in advance.

Analysis Procedure

1. Click the Phylogenetic Tree (using profile) icon from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.
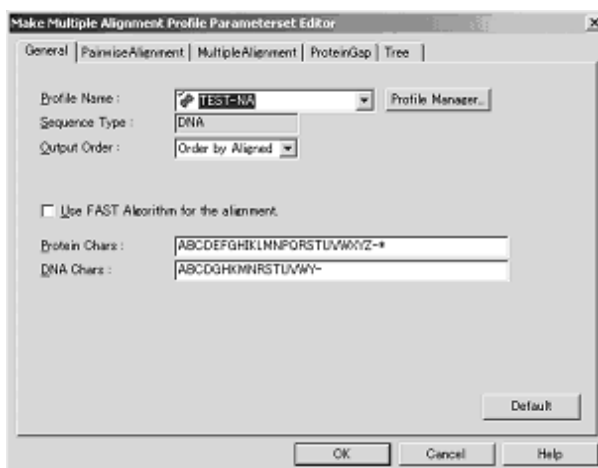
2. In the Profile Name field, select a file you want to create, and then click the OK button.
3. In the Sequence View, select a sequence you want to analyze (a sequence to be added to a tree) as the target.
4. Click the Phylogenic Tree (using profile) button in the Analysis menu.

Explanation of the Result Window

When a branch is added to a phylogenic tree created using a profile, the branch is shown red.   For details about the window that displays phylogenic trees, refer to "Explanation of the Result Window" in "3.23 Phylogenic Tree-DNA".
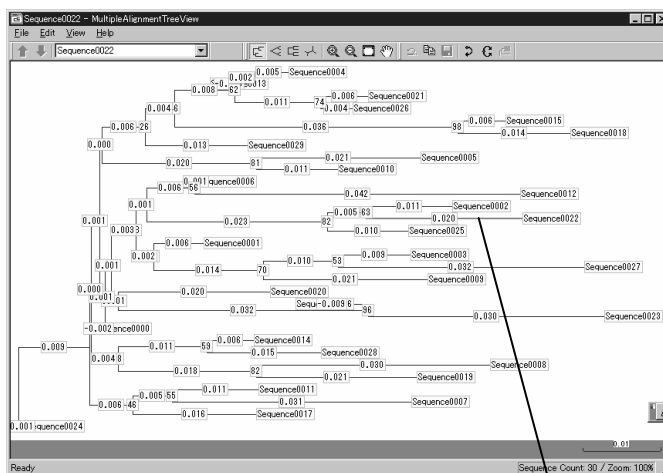
Added sequence

3.27 Sequence Assembly

After importing sequences to be assembled or clustered into the project, launch DNASIS Assemble from Analysis tools menu under DNA Multiple Sequence..

Explanation of the Result Window
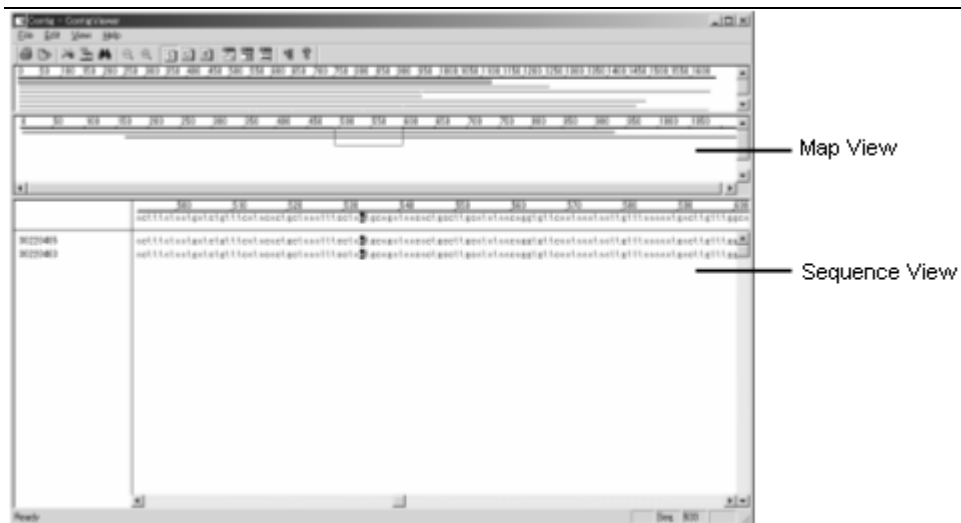
Map View

Display the contig and fragments graphically.

Sequence View

Display the contig and fragments.

Setting Parameters

1. Click the Sequence Assemble icon from analysis button view. An Analysis dialog box will appear then click the Parameter button and the Sequence Assemble Parameterset Editor will appear.
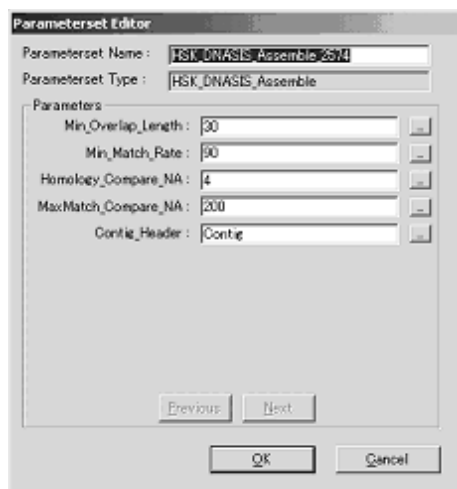
2. Set each parameter.
3. Click the OK button.

3.28 Clustering

This function sorts all the sequences displayed in the window into some clusters according to the similarity of sequences.   From the list, you can identify the cluster to which each sequence belongs.   The tool is launched from the Analysis tools menu under DNA Compare.

Explanation of the Result Window

From left to right, each cell shows the following: the input sequence name, the cluster name to which the sequence belongs, the homology score with the cluster-representing sequence, and the input data sequence.

The longest one of the sequences under the same cluster is chosen as the cluster-representing sequence.
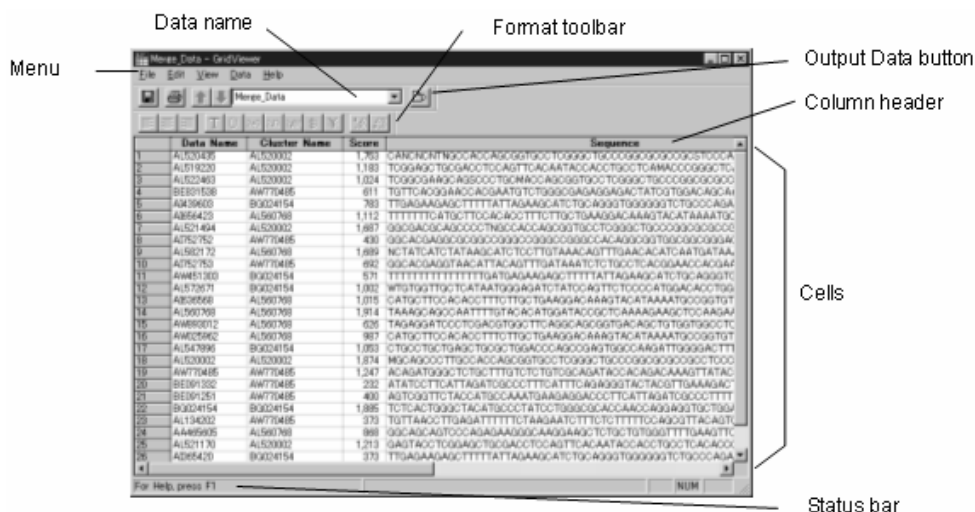
Cells

The following shows how to select cells.   (It is the same method as how to operate the Excel program.)

Select a particular column: Click the column number.

Select a particular row: Click the row number.

Select all cells: Click the Select All Cells button.

Select a range of neighboring cells:

Click the upper left cell of a selection range.   Then, while holding down the Shift key, click the lower right cell.

Select a range of non-neighboring cells:

Click the first cell.   Then, while holding down the Ctrl key, click the subsequent cells.

Each time you double-click the column header, the cells are sorted in ascending or descending order.   Note that you can edit cells but cannot save them.

Output Data Button

The button is not used for this analysis.

Grid Viewer Menu

File menu	Description
Export...	Outputs the entire data into a text file.
Print Preview	Displays a print preview.
Print Setup...	Makes a printer setting.
Print...	Starts printing.
Exit	Closes the window.

Edit menu	Description
Undo	Cancels the previous operation.
Cut	Cuts the data.
Copy	Copies the data.
Paste	Pastes the data.
Select All	Selects everything.
Find...	Attempts to find the target.
Find Again	Attempts to find the next target.

View menu	Description
Navigation Toolbar	Toggles the Navigation toolbar to display or hide it.
Format Toolbar	Toggles the Format toolbar to display or hide it.
Status Bar	Toggles the status bar to display or hide it.

Help menu	Description
Contents	Displays online help.
About GridViewer...	Displays the version information about GridViewer.

Setting the Clustering Standard

1. Click the Clustering button from analysis button view and an Analysis dialog box will appear.   Click the Parameter button and a Sequence Clustering Parameterset Editor will appear.
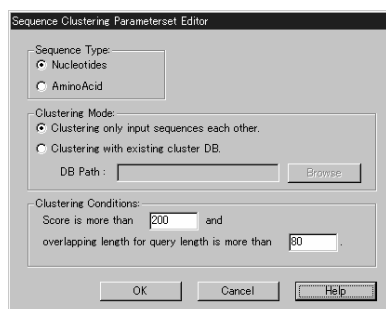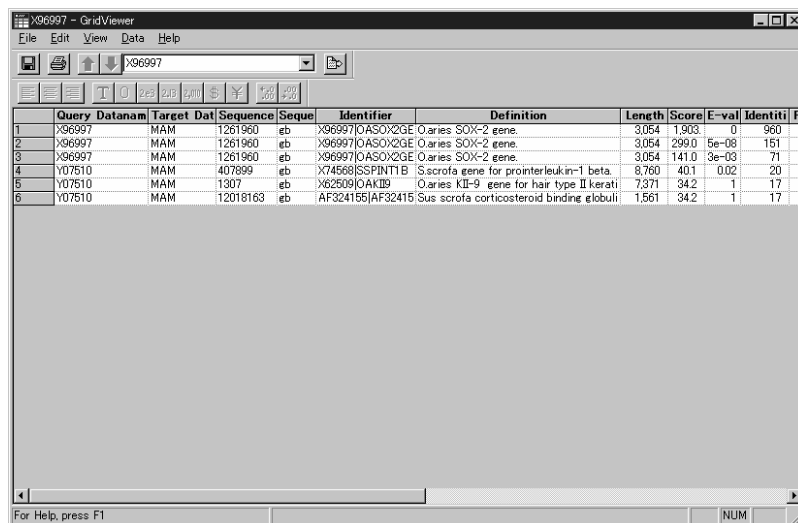
2. In the Clustering Conditions field, specify the score and the overlap length.   Those sequences meeting these similarity standards are classified under the same cluster.

3. Click the OK button.

3.29 BLAST Search and Extraction (Report)

This function performs BLAST-search for all the DNA sequences displayed in the window, takes out those sequences with higher similarity, and produces a list containing the results.

Explanation of the Result Window

File menu	Description
Export...	Outputs the entire data into a text file.
Print Preview	Displays a print preview.
Page Setup...	Makes a printer setting.
Print...	Starts printing.
Exit	Closes the window.

Edit menu	Description
Undo	Cancels the previous operation.
Cut	Cuts the data.
Copy	Copies the data.
Paste	Pastes the data.
Select All	Selects everything.
Find...	Attempts to find the target.
Find Again	Attempts to find the next target.

View menu	Description
Navigation Toolbar	Toggles the Navigation toolbar to display or hide it.
Format Toolbar	Toggles the Format toolbar to display or hide it.
Status Bar	Toggles the status bar to display or hide it.

Data menu	Description
Previous Data	Displays the previous data item when multiple items are opened at the same time.
Next Data	Displays the subsequent data item when multiple items are opened at the same time.

Data menu	Description
Sort	Sorts all data item in ascending or descending order.

Help menu	Description
Contents	Opens the Help.
About GridViewer...	Displays the version.

Button		Description
	Export button	The same function as the Export menu.
	Print button	The same function as the Print menu.

Cells

Query Dataname	Shows the name of a query sequence.
Target Database	Shows the database that is the target of BLAST search.
Sequence ID	Shows the ID of the entry in the original database where the subject sequence has been registered.
Sequence Type	Shows the original database where the subject sequence has been registered. Example: gb: GenBank, emb: EMBL, dbj: DDBJ, etc.
Identifier	Shows the identifier of the subject sequence.
Definition	Shows the definition of the subject sequence.
Length	Shows the length of the subject sequence.
Score	Shows the score of a match. A match with a higher score value is higher in similarity.
E-value	Shows the expected value of a match. A match with a lower score value is higher in similarity.
Identities	Shows the percentage of the matching bases (or amino acids) within the entire length of a match.
Positives	Shows the number of groups in which the score has a positive value within the entire length of a match, when the query sequence and the subject sequence are compared for each amino acid.
Overlap Length	Shows the length of a match.
Gaps	Shows the total number of gaps inserted into the query sequence and the subject sequence. This cell remains blank when there is no gap.
Strand	Shows the direction of the match (for example, from 3' to 5' or from 5' to 3').
Matching Percentage	Shows the matching rate.
Query Start	Shows the start point of a match in a query sequence.
Query End	Shows the end point of a match in a query sequence.
Target Start	Shows the start point of a match in a subject sequence.
Target End	Shows the end point of a match in a subject sequence.
Query Length	Shows the length of a query sequence.
Query Identifier	Shows the identifier of a query sequence.

Specifying a Database to Be Searched

Select a database as the target of homology search.   You can select more than one database at one time.

1. Click the BLAST Search & Report button under DNA Compare in Analysis Tools menu.   An Analysis dialog box will appear.   Click the Parameter button and Analysis Parameters will appear.

2. Select BLAST Search and click Set... to display BLAST Parameters.

3. In the Nucleotide Database field, place a checkmark in the check box for the target database.

4. Click the OK button to complete the setting.

Setting Extract Conditions

1. Click the BLAST Search & Report button and an Analysis dialog box will appear.   Click the Parameter button and Analysis Parameter will appear.

2. Select Make Report and click Set....   The Collect Homology Results Parameter Editor window appears.

*Refer to "Collect Homology Results Parameter Editor" in "4.27 Blast Search and Extraction".
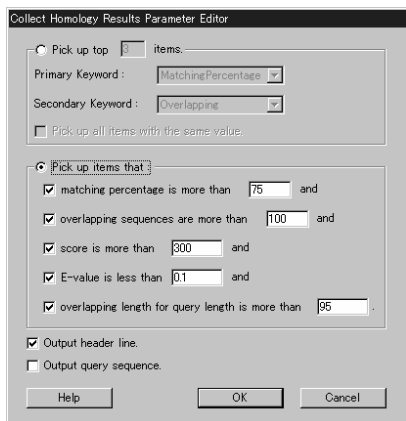
3. Set the extract conditions*.

4. Click the OK button to complete the setting.

3.30 Amino Acid Content

This function analyzes amino acid sequences and displays the result of analyzing the amino acid content.

Explanation of the Result Window
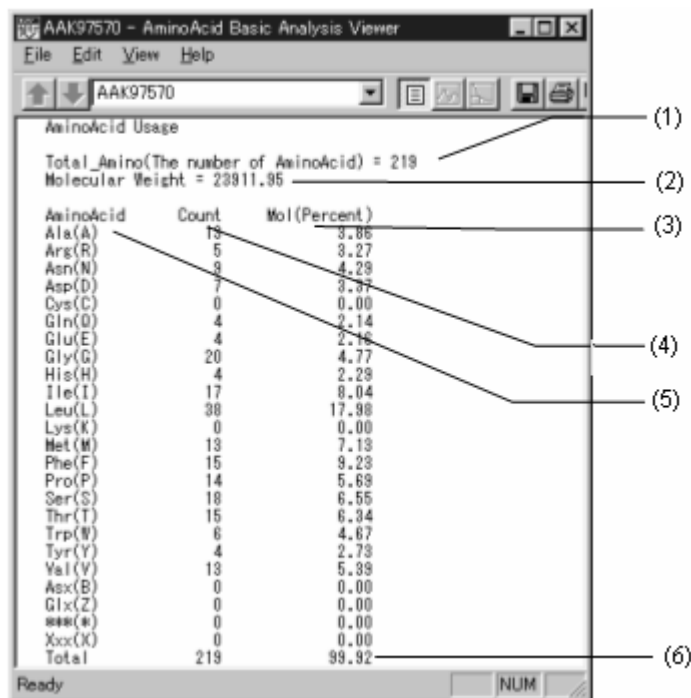
(1)	Total number of amino acid residues
(2)	Total molecular weight
(3)	Molar ratio
(4)	Number of amino acid residues
(5)	Amino acid name
(6)	Total molar ratio

File menu	Description
Export	Exports the data in the window into a text file.
Print	Prints the window.
Print Preview	Displays an on-screen preview of the printed image.
Page Setup	Provides various print settings.
Exit	Terminates View.

Edit menu	Description
Copy	Copies the data in the window as a tabbed character string into the clipboard.

View menu	Description
Toolbar	Toggles the toolbar to display or hide it.
Status bar	Toggles the status bar to display or hide it.

Help menu	Description
About DNABasicAnalysisViewer	Displays the version information for this View in the dialog box.
Contents	Displays online help.

Button		Description
	Export button	The same function as the Export menu.
	Print button	The same function as the Print menu.
	Copy button	The same function as the Copy menu.

3.31 Isoelectric Points

This function analyzes amino acid sequences and displays the result of analyzing isoelectric points.

Explanation of the Result Window
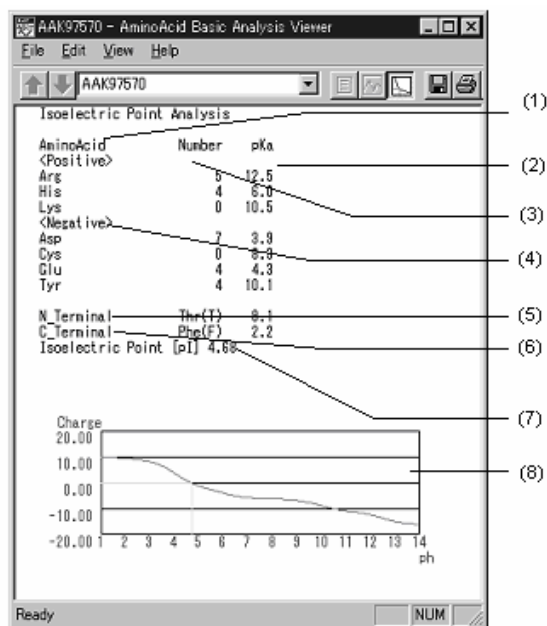
(1)	Amino acid name having positive charge
(2)	Amino acid name having negative charge
(3)	Charge weight
(4)	Number of amino acid residues
(5)	Residue at N end
(6)	Residue at C end
(7)	Isometric point
(8)	Charge and pH graph

File menu	Description
Export	Exports the data in the window into a text file.
Print	Prints the window.
Print Preview	Displays a printing image.
Page Setup	Provides various print settings.
Exit	Closes the window.

Edit menu	Description
Copy	Copies the data in the window as a tabbed character string into the clipboard.

View menu	Description
Toolbar	Toggles the toolbar to display/hide it.
Status bar	Toggles the status bar to display/hide it.

Help menu	Description
About DNABasicAnalysisViewer	Displays the version information for View in the dialog box.
Contents	Displays online help.

Button		Description
	Export button	Exports the data in the window into a text file. It is possible to export data for each part.
	Print button	Prints the window.
	Copy button	Copies the data in the window as a tabbed character string into the clipboard.
	Horizontal View Expansion button	Expands the view horizontally.
	Horizontal View Shrinkage button	Shrinks the view horizontally.
	Vertical View Expansion button	Expands the view vertically.
	Vertical View Shrinkage button	Shrinks the view vertically.
	Help button	Displays online help.

3.32 Hydrophilicity, Hydrophobicity, and Secondary Structure

This function analyzes the hydrophilicity, hydrophobicity, and secondary structure for an amino acid sequence using parameters pertaining to the hydrophilicity, hydrophobicity, and secondary structure, and then displays the results graphically.

Explanation of the Result Window
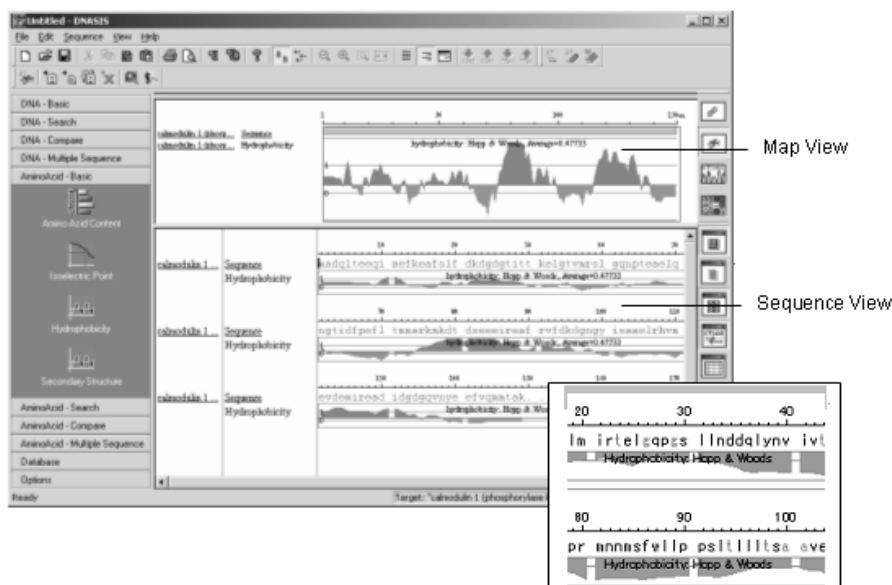
Map View

This view displays the entire sequence graphically.

Sequence View

This view displays the result of analysis for the specified table graphically.   The table name and average value are also shown at the center of the graph.

Selecting a Table
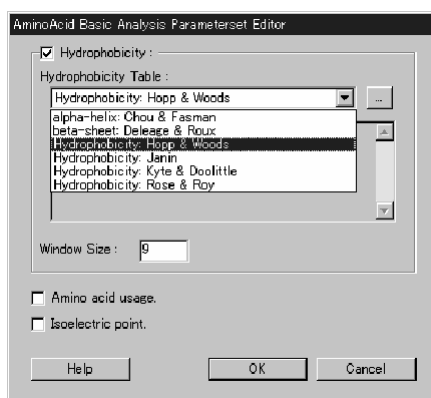
1. Click the Hydrophilic/Hydrophobic Search button from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.

2. In the Hydrophobicity Table field, select a table you want to use.   The description of each table is then shown.

3. Click the OK button.

Creating and Editing a New Table

1. Click the Hydrophilic/Hydrophobic Search button from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.

2. Click the ... button in the Hydrophobicity Table field. Window appears.

3. Click the New... button to display the New Hydrophobicity Table, as shown in the figure.
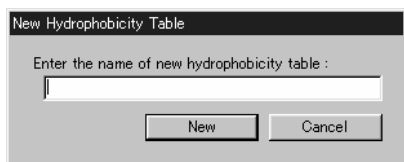
4. Enter the name of a table you want to create and click the New button.   The display returns to the Hydrophobicity Table Editor window.

5. In the Hydrophobicity Table Editor window, edit the contents of the table you created.

6. Click the OK button to return to Amino Acid Basic Analysis Parameterset Editor.

7. Click the OK button.

3.33 Motif Search - Amino Acid

This function searches for the motif of data about amino acid sequences.   There are two available methods: one uses a database and the other searches for any pattern you have entered.

Explanation of the Result Window
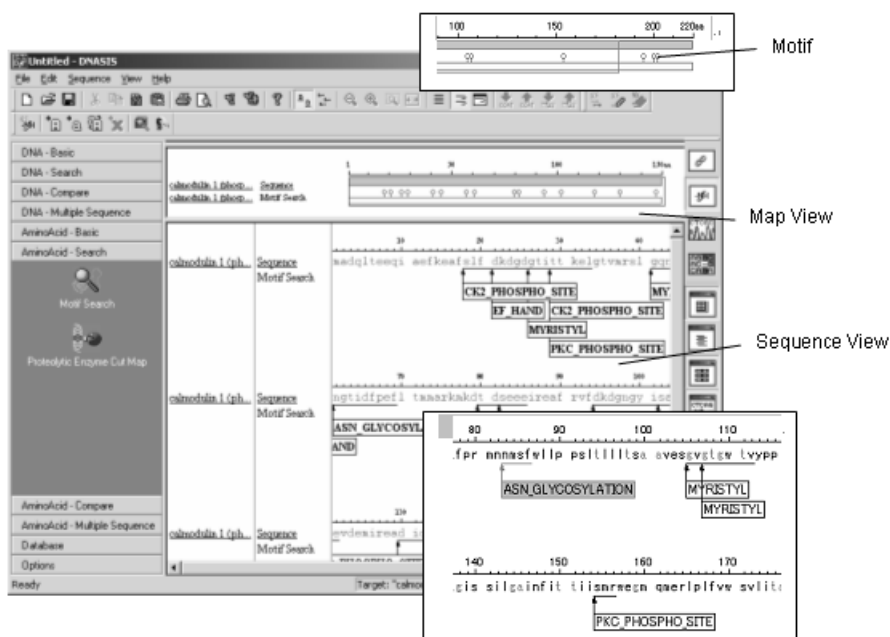
Map View

A pin shows the retrieved motif.   If you move the cursor to the pin and click the mouse, the display color changes to the selecting color, indicating the motif is selected.   At the same time, the sequence in the motif region is also selected.

Sequence View

Together with the sequence, this View displays: the motif name and the identified part.   Click the mouse on a motif to select it.   At the same time, the sequence in the motif region is also selected.   What is displayed within a red frame in the Map View is now displayed in the Sequence View.

Search Using a Motif Database

1. Click the Motif Search - Amino Acid button from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.
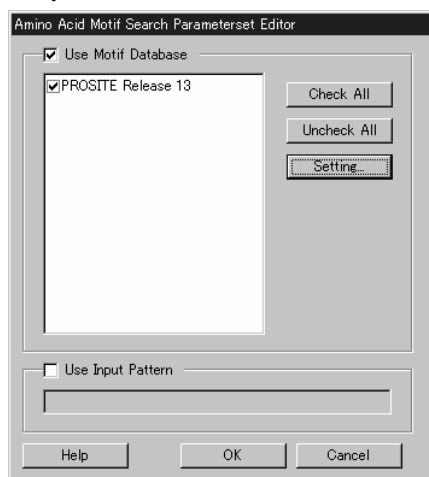
2. Place a checkmark in the Use Motif Database check box and select an appropriate database from the list of databases displayed.   To create a new database, click the Setting... button and use Amino Acid Motif Database Manager.

3. Click the OK button, to complete the setting.

Search by Entering a Motif Pattern

1. Click the Motif Search/Amino Acid button from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.
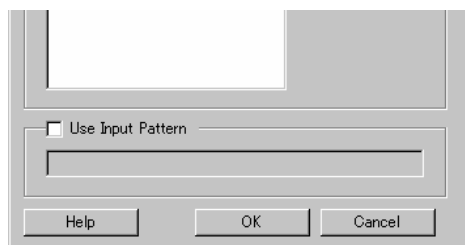
2. Place a checkmark in the Use Input Pattern check box and enter a pattern you want to search for.

3. After selecting Use Input pattern and entering desired pattern in the text box, click the OK button.

4. Analysis Result View shows a motif, if found, with the name of Input_Pattern.

Creating a Motif Database

1. Click the Motif Search/Amino Acid button from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.

2. Click the Setting... button.   Amino Acid Motif Database Manager appears.

3. Click the New button.   A database named "Untitled" is created in the window.

4. Double-click "Untitled".   The Database Property window appears.
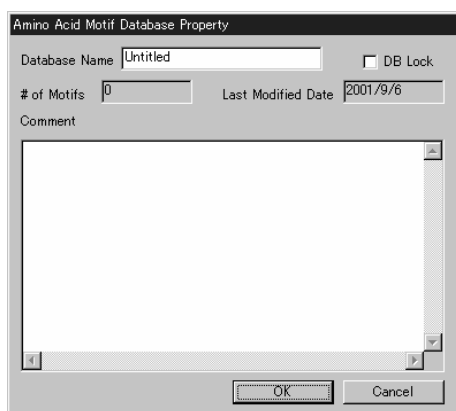
5. Make settings according to the following contents of the window.

Database Name:	Name of a database to be created
DB Lock:	When checked, this item prevents motifs from being added or deleted or prevents a motif database from being deleted.
# of Motifs:	Number of motifs registered with the database
Last Modified Date:	Date on which data was last modified
Comment:	Comment given to a database

6. When you complete the selection, click the OK button.   This concludes the process of creating the motif database.

Adding Motif Data

1. Click the Motif Search/Amino Acid button from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.

2. Click the Setting... button.   Amino Acid Motif Database Manager appears.

3. Select a motif database to which you want to add motif data.

4. Click the View button to display a list of motifs registered with the database.

5. Click the New button to display the Amino Acid Motif Property dialog box, as shown in the figure.
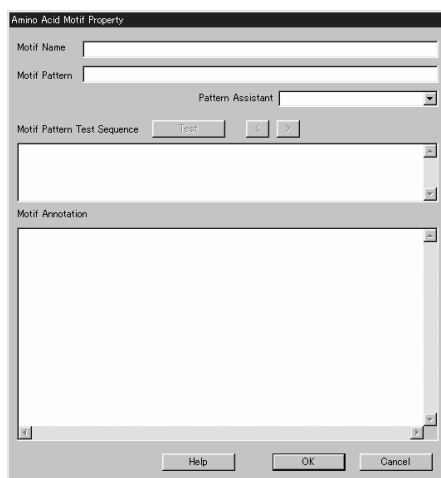
6. Enter the motif name, motif pattern, and motif annotation (optional), and then click the OK button.

7. The new motif is added to the list.

8. To edit motif data, select the motif and click the Property button.

Browsing the Detail of a Motif Searched for

If you double-click a motif in the Sequence View, the details of the motif are displayed.
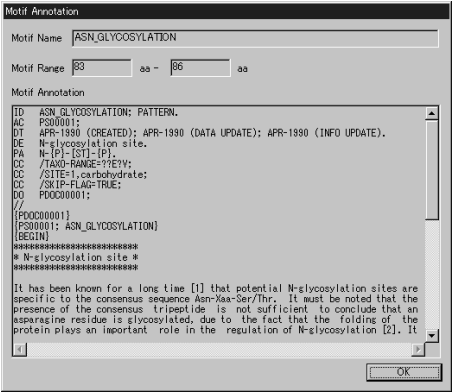
Displaying a List of Search Results

For the motif result in sequence view, select the sequence name and analysis name then click the Result List Dialog button.

This displays a list of motifs retrieved.
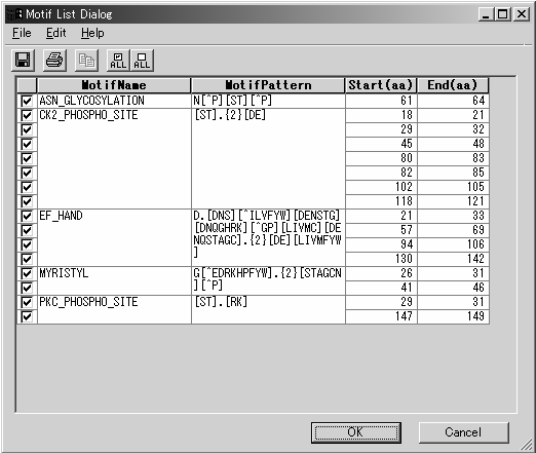
3.34 Common Motif Search

Analyzes motifs common to multiple sequences. Searches can be done using the database or by specified patterns.

Common motifs to either DNA sequences or amino acid sequences can also be searched.

Result Window Description
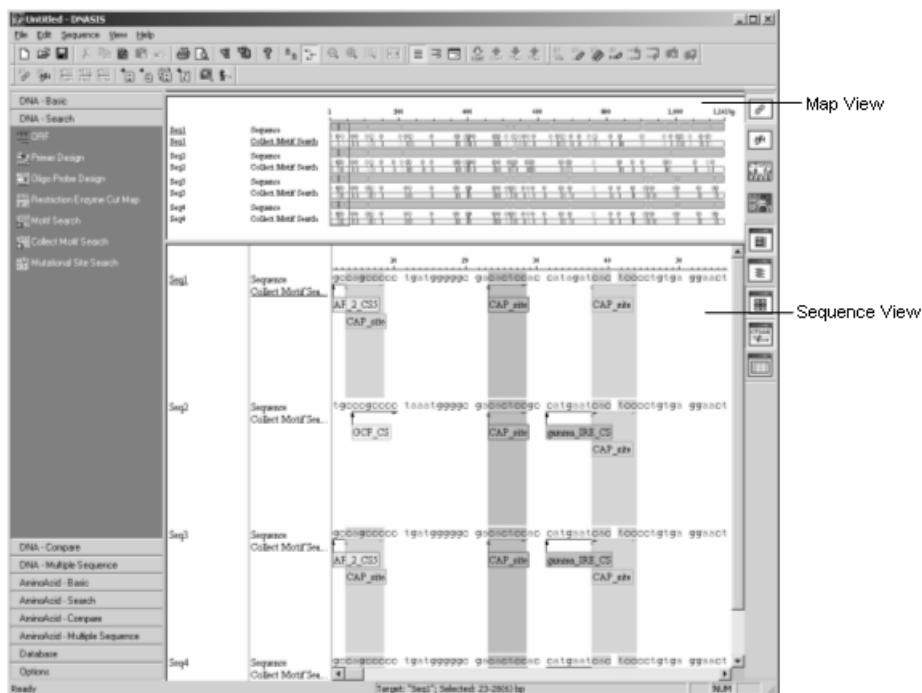
Map View

Displays the common motifs of the pin search. Align the cursor to the pin and click to color and highlight it as selected.

Sequence View

Displays the sequence together with the common motif name and recognized parts. Click the motif name to color and highlight it as selected.

Search with the Motif Database (DNA)

1. Click the Common Motif Search icon from analysis button view and an Analysis dialog box will appear. Then click the Parameter button. The dialog below appears.
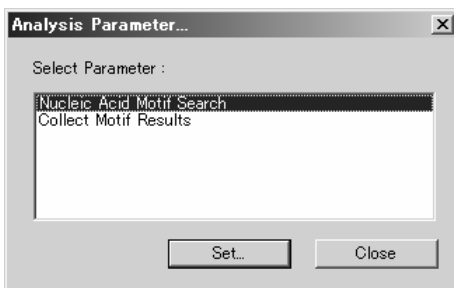
2. Select the Nucleic Acid Motif Search under Select Parameter then click Set. The Nucleic Acid Motif Search Parameter Set Editor appears.

3. Select Use Motif Database then select the database from the list.
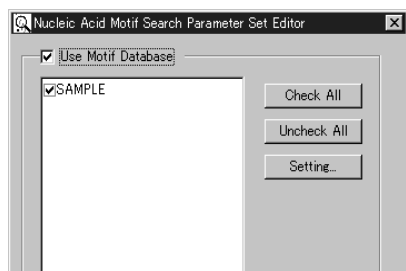
4. Click OK.

Search by entering the Motif Pattern (DNA)

1. Click the Common Motif Search icon from analysis button view and an Analysis dialog box will appear. Then click the Parameter button. The dialog below appears.
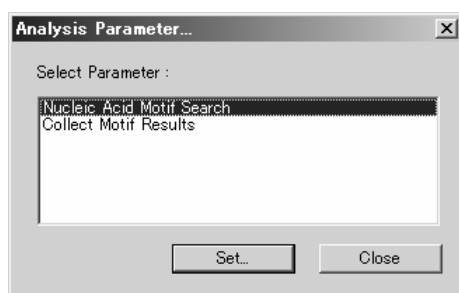
2. Select Nucleic Acid Motif Search under Select Parameter then click Set. The Nucleic Acid Motif Search Parameter Set Editor appears.

3. Select Use Input Pattern then enter or paste the motif to search.
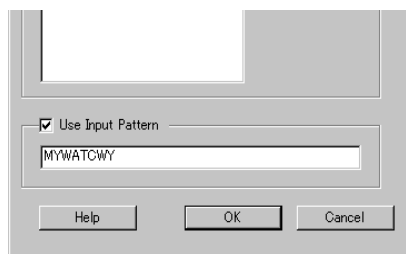
4. Click OK.

Search with the Motif Database (Amino Acid)

1. Click the Common Motif Search icon from analysis button view and an Analysis dialog box will appear. Then click the Parameter button. The dialog below appears.
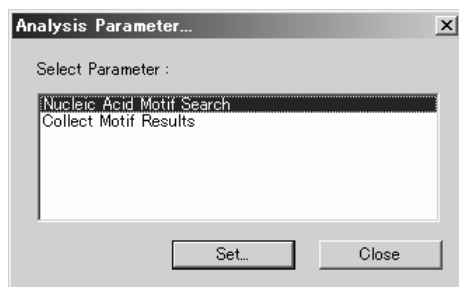
2. Select Nucleic Acid Motif Search under Select Parameter then click Set. The Amino Acid Motif Search
   Parameter Set Editor appears.

3. Select Use Motif Database then select the database from the list.
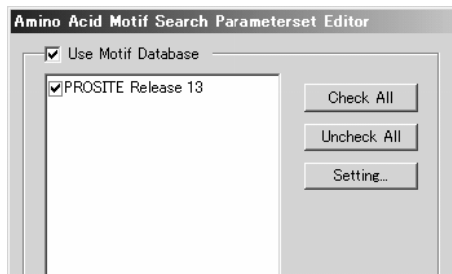
4. Click OK.

Search by entering the Pattern (Amino Acid)

1. Click the Common Motif Search icon from analysis button view and an Analysis dialog box will appear. Then
   click the Parameter button. The dialog below appears.
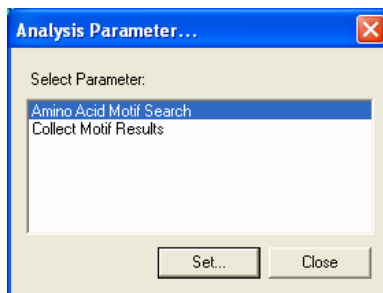
2. Select Nucleic Acid Motif Search under Select Parameter then click Set. The Nucleic Acid Motif Search
   Parameter Set Editor appears.

3. Select Use Input Pattern, then input or paste the motif to search.
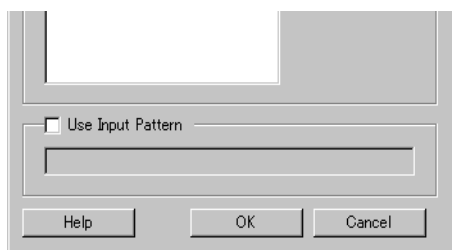
4. Click OK.

Setting the Search Method

1. Click the Common Motif Search icon from analysis button view and an Analysis dialog box will appear. Then
   click the Parameter button. The dialog below appears.
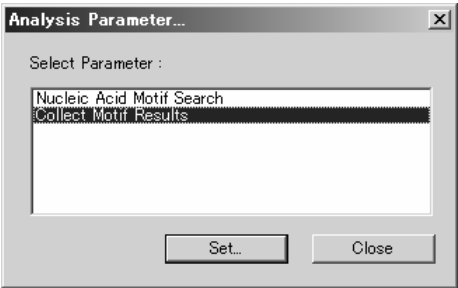
2. Select Collect Motif Results then click Set. The dialog below appears.
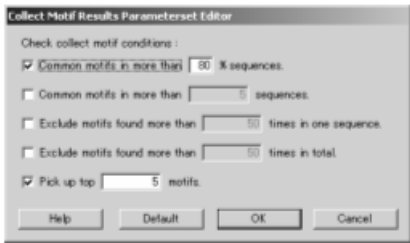
Item	Description
Common motifs in more than X % sequences.	Select to designate as common motifs when they are common in more than the specified percent sequences for motif search results input simultaneously.
Common motifs in more than X sequence.	Select to designate as common motifs when they are common in more than specified number of sequences for motif search results input simultaneously.
Exclude motifs found more than X times in one sequence.	Select to exclude motifs found more than the specified number of times in a certain sequence.
Exclude motifs found more than X times in total..	Select to exclude motifs found more than the specified number of times in all the sequences.
Pick up top X motifs.	Select to designate specified number of common motifs counting from the largest number of motifs found. Motifs with the same number are all regarded as common.

3. Select the parameter then click OK.

List up Search Results

Select the Common Motif Sequence, right click and select the Show Common Motif Dialog, or for a frame in sequence view select the sequence name and analysis name then click the Result List Dialog button, and the search result list will appear.
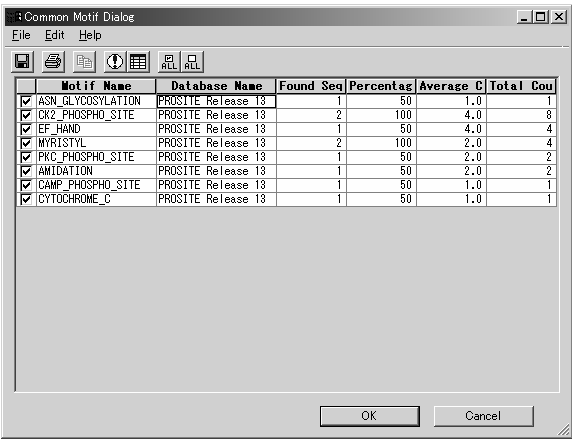
Use Copy All or Copy Selected from Edit in the menu to copy all or selected cells to the clipboard as tab delimited text.

Use Save All as or Save Selected as from File in the menu to store all or selected cells in a file as tab delimited text.

Browsing Annotations of Searched Common Motifs

Click ![icon] from the Common Motif dialog to display a motif annotation.
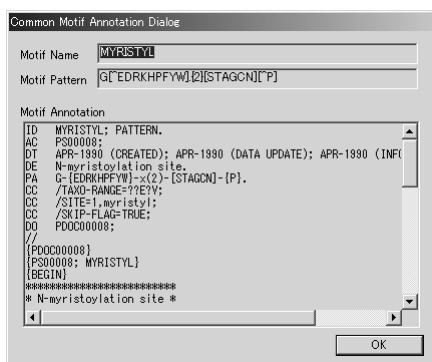
Browsing Details of Searched Common Motifs

Click ![icon] from the Common Motif dialog to display motif details.
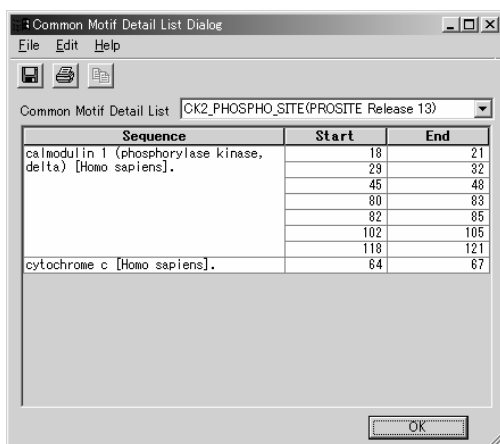
Use Copy All or Copy Selected from Edit in the menu to copy all or selected cells to the clipboard as tab delimited text.
Use Save All as or Save Selected as from File in the menu to store all or selected cells in a file as tab delimited text.

3.35 Proteolytic Site Search

This function searches through the amino acid sequence to identify the areas split by the proteolytic enzymes and displays the result of search.

Explanation of the Result Window

Map View
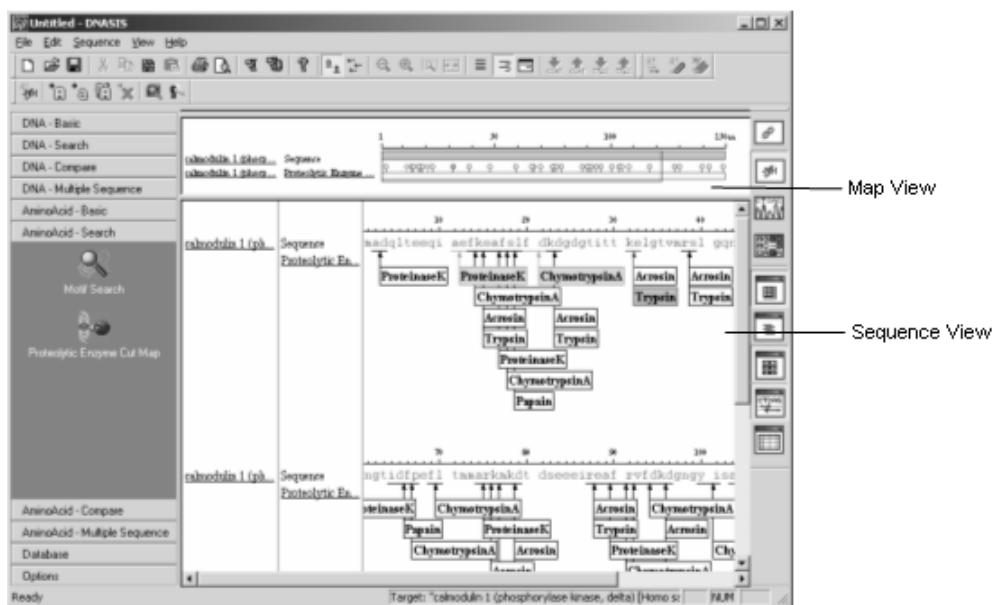
Sequence View

Map View
A pin shows the identified split area for the proteolytic enzymes.   If you move the cursor to the pin and click the mouse, the display color changes to the selection color, indicating the area is selected.

Sequence View
Together with the sequence, this view displays the proteolytic enzyme name and the proteolytic site.   Clicking the mouse on a proteolytic enzyme turns it to the selection color.

Selecting Proteolytic Enzymes to Be Searched for

Proteolytic enzyme is registered with the Proteolytic enzyme database.   Only the enzymes selected from the database are searched for.   The more enzymes to be searched for, the longer it takes to perform the search and display its result. It is recommended that you only select Proteolytic enzymes you want to search for before starting actual search.   Select the Proteolytic enzymes according to the procedure.

1. Click the Proteolytic Recognition Site Search icon and an Analysis dialog box will appear. Click the Parameter button

   and a Proteolytic Site Parameter Editor will appear.
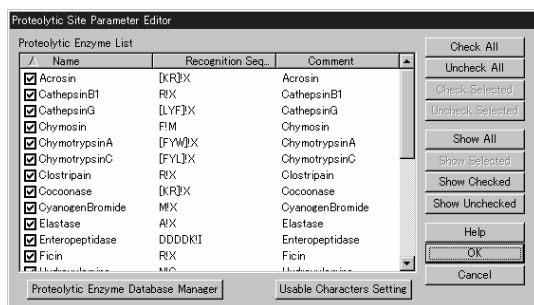
2. Proteolytic Site Parameter Editor displays the proteolytic enzymes registered. A restriction enzyme is selected when

   the preceding check box is marked with a checkmark.   Place a checkmark on the check box for the restriction enzyme

   you want to search for.

3. Click the OK button.

Registering a New Proteolytic Enzyme

1. Click the Proteolytic Recognition Site Search icon from analysis button view and an Analysis dialog box will appear.
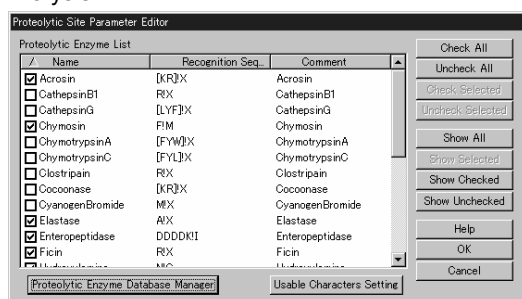
2. Click the Proteolytic Enzyme Database Manager button at the bottom of the window.

3. Click New on the Proteolytic Enzyme Database Manager window to display the New Database dialog box.

4. Enter the enzyme name you want to register, the sequence, and the comment (optional), and then click the OK button.

Displaying a List of Split Areas by Proteolytic Enzymes

From analysis result in sequence view, select the sequence name and analysis name then click the Result List Dialog button.   The list of split areas appears.
You can copy and save any data in the window.
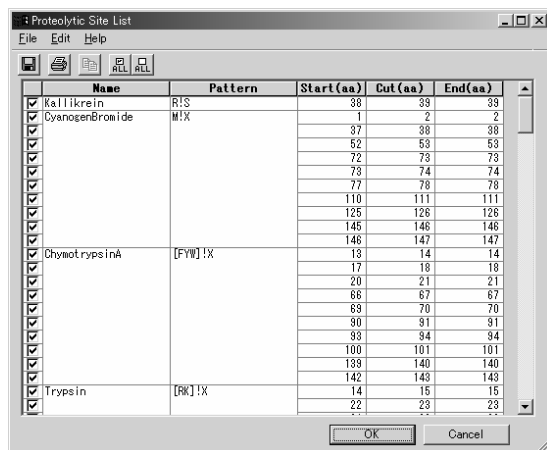
Use Copy All or Copy Selected Cells from Edit in the menu to copy all or selected cells to the clipboard as tab delimited text.
Use Save All as or Save Selected as from File in the menu to store all or selected cells in a file as tab delimited text.

Selecting a Proteolytic Enzyme to Be Displayed

1. Right-click the result of analysis in the Sequence View and select the Proteolytic Site List.   The list of split areas is displayed.

2. Click the check box at the leftmost of the list to select any item or items you want to display.   Uncheck those items you want to hide.

3. Click the OK button to complete the setting.

3.36 BLAST Search (Amino Acid)

Types of BLAST Search

There are two ways of BLAST search for amino acid sequences.
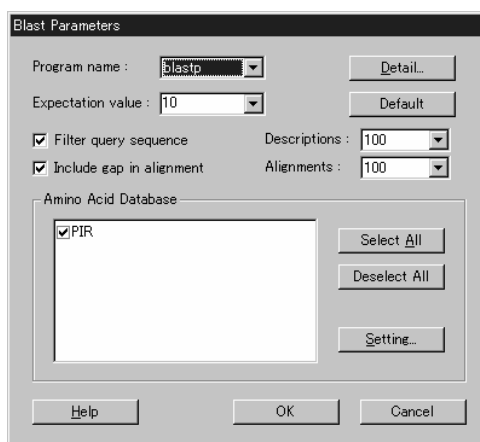
Button name	Program name	Description
BLAST search	blastp	Performs homology search between amino acid sequences and an amino acid database.
BLAST search (Translation DB)	tblastn	Performs homology search between amino acid sequences and an amino acid database translated in all frames.
One-to-One BLAST Search	blastp	Performs a one-to-one BLAST search between two amino acid sequences.

Explanation of the Result Window

Refer to the Explanation of the Result Screen in Section 19, "BLAST Search".

Selecting a Database to Be Searched

1. Click the BLAST Search icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and BLAST Parameters will appear.

2. The Amino Acid Database field lists the databases.   Place a checkmark in the check box for the target database.

3. Click the OK button.

3.37 Internet BLAST Search (Amino Acid)

Types of BLAST Search

There are three ways of BLAST search for amino acid sequences.

Button name	Program name	Description
BLAST search	blastp	Performs homology search between amino acid sequences and an amino acid database.
BLAST search (Translation DB)	tblastn	Performs homology search between amino acid sequences and an amino acid database translated in all frames.
One-to-One BLAST Search	blastp	Performs a homology search between two different amino acids.

Explanation of the Result Window

Refer to the Explanation of the Result Window in Section 19, "BLAST Search".

Selecting a Database to Be Searched (excluding one-to-one BLAST search)

1. Click the BLAST Search button from analysis button view and an Analysis dialog box will appear. Then click the Parameter button.

2. Click the Setting... button to display the NCBI Advanced BLAST Search window.

3. From the Database Selection combo box, select a database you want to search through.

4. Click the OK button.

3.38 Smith-Waterman Search (Amino Acid)

This function provides high-precision homology search using the Smith-Waterman algorithm between the input sequence and the target database and which prevents search misses occurring in the FASTA or BLAST algorithm.

Types of Smith-Waterman Search

The Smith-Waterman search has two types for an amino acid sequence.

Button name	Description
Smith-Waterman search	Performs a Smith-Waterman search between an amino acid sequence and an amino acid sequence database.
One-to-One Smith-Waterman Search	Performs a Smith-Waterman search between two different amino acid sequences.

Explanation of the Result Window

Refer to the Explanation of the Result Window in Section 19, "BLAST Search".

Selecting a Database to Be Searched (Smith-Waterman search only)

1. Click the Smith-Waterman Search icon from analysis button view and an Analysis dialog box will appear. Click the Parameter button and a GENE BRIGHT III Parameterset Editor will appear.

2. In the Target Database field, place a checkmark in the check box for the database you want to search through.

3. Click the OK button.

3.39 Multiple Alignment (Amino Acid)

This function performs multiple alignment, or optimum alignment of multiple sequences, using multiple sequences that have been entered by the Editor.   The algorithm used here is Clustal W.

Explanation of the Result Window

Refer to the Explanation of the Result Window in Section 22, "Multiple Alignment".

Setting Criteria for Determining Match Bases

The result of multiple alignment is color-coded according to the match rate.
You can change the match rate and color combination.

1. Click the Parameters icon on the toolbar to open Sequence Editor Parameter Set Editor.

2. Click the Sequence tab.
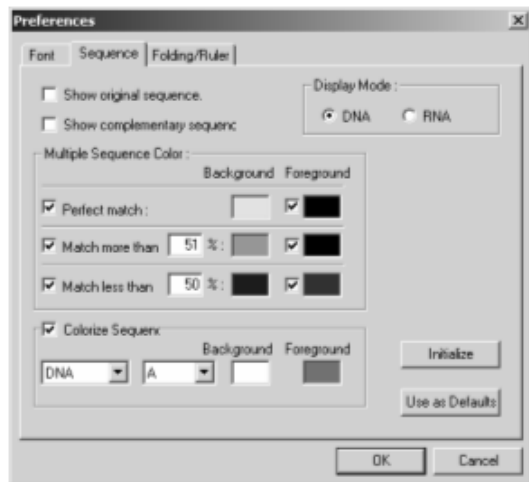
3. The number in the Multiple Sequence Color indicates the matching rate.   To its right, the display color is shown.

   Additionally, in order to enable this setting the next time the program starts up, Click Use Default. Also, click

   Initialize to restore the factory setting.

Analyzing a Selected Range

In the alignment display mode, you cannot perform other types of analysis unless you cancel this.

1. Select a region you want to analyze, as shown in the figure.

2. On the View Toolbar, click the Alignment icon to cancel the alignment display mode.   Now you can move on to

   analysis.

3. Start the analysis.   Because the range selected in step 1 above is interlocked here, it gives a rough measure of the

   region of analysis.

Creating a Consensus Sequence

Based on the result of alignment, the most frequently occurring amino acid is selected as the consensus for each position.
From menu bar in the Alignment Mode Window, Select Sequence and then Make Consensus. The consensus sequence is added to the Sequence View.

```
                                   10        20        30        40        50
M0001        Sequence     TGCCCGCCCC TAAATGGGG- CGACACTCCG CCATGAATCA CTCCCCTGTG AG
M0002        Sequence     -GCCAGCCCC CTGATGGGGG CGACACTCCA CCATAGAATCA CTCCCCTGTG AG
M0003        Sequence     -GCCAGCCCC CTGATGGGGG CGACACTCCA CCATGAATCA CTCCCCTGTG AG
M0004        Sequence     -GCCAGCCCC CTGATGGGGG CGACACTCCA CCATGAATCA CTCCCCTGTG AG
Untitled001  Consensus    TGCCAGCCCC CTGATGGGGG CGACACTCCA CCATGAATCA CTCCCCTGTG AG

                                   70        80        90        100       110
M0001        Sequence     GTCTTCACGC AGAAAGCGTC TAGCCATGGC GTTAGTATGA GTGTCGTACA GC
M0002        Sequence     GTCTTCACGC AGAAAGCGTC TAGCCATGGC GTTAGTATGA GTGTCGTCCA GC
M0003        Sequence     GTCTTCACGC AGAAAGCGTC TAGCCATGGC GTTAGTATGA GTGTCGTCCA GC
M0004        Sequence     GTCTTCACGC AGAAAGCGTC TAGCCATGGC GTTAGTATGA GTGTCGTCCA GC
Untitled001  Consensus    GTCTTCACGC AGAAAGCGTC TAGCCATGGC GTTAGTATGA GTGTCGTCCA GC

                                   130       140       150       160       170
M0001        Sequence     CCCCCCCTCC CGGGAGAGCC ATAGTGGTCT GCGGAACCGG TGAGTACACC GG
M0002        Sequence     CCCCCCCTCC CGGGAGAGCC ATAGTGGTCT GCGGAACCGG TGAGTACACC GG
```

3.40 Phylogenic Tree (Amino Acid)

This function calculates the phylogenic tree by using an input of three or more amino acid sequences and displays the result of calculation.

Explanation of the Result Window

Refer to the Explanation of the Result Window in Section 23, "Phylogenic Tree-DNA".

Changing the Type of a Phylogenic Tree

You can select a phylogenic tree from four types: Phylogram, Slanted cladogram, Rectangular cladogram, and Unrooted.   From the Tree View toolbar, select any type you want to display.
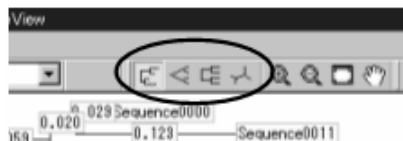
Changing the Font

1. Select View-Preferences... from the menu bar to display Parameter Set Editor.
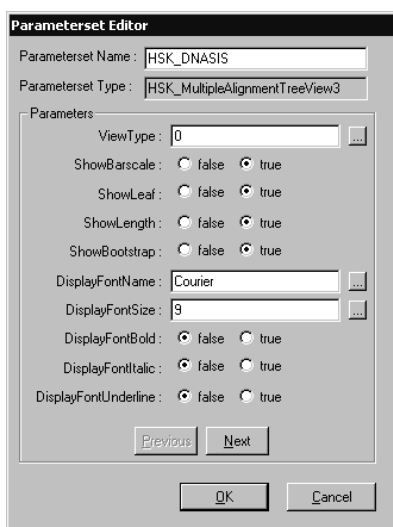
2. To change the font, use the Display Font Name field.   To change the size, use the Display Font Size field.

3. Click the OK button to complete the setting.   The phylogenic tree under a new setting appears.

Displaying an Expanded Phylogenic Tree

1. Click the 🔍 icon on the toolbar.   The shape of the mouse cursor turns to a magnifying glass.

2. Click or drag any portion you want to expand.   The specified portion is expanded.

   To shrink it, click the 🔍 button and perform a similar operation.   To return the display to its original size, click the ⬜ button.

Setting an Out-Group

You can set the selected branch as an out-group.

1. Click the ↻ icon on the toolbar.   The shape of the mouse cursor changes to the + mark.

2. Move the cursor to a branch you want to set as an out-group and click it.   The specified branch is now set in the out-group.

Replacing Branches

You can replace branches.

1. Click the ![icon] icon on the toolbar.   The shape of the mouse cursor changes to the + mark.

2. Move the cursor to a branch you want to replace with another within a tree and click it.   The specified branch is now
   replaced and displayed.

Evaluating the Branching Reliability (Bootstrap Tree)

This function evaluates the reliability of a tree form using the bootstrap method.

1. Click the Phylogenetic Tree icon from analysis button view and an Analysis dialog box will appear. Click the
   Parameter button and a Multiple Alignment Parameter Editor will appear.
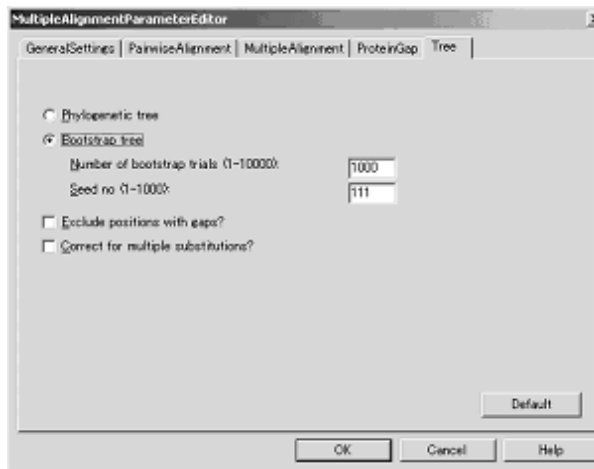
2. Click the Tree tab.

3. Select Bootstrap tree.

Number of bootstrap trials:	The number of random numbers that occurred
Seed No:	The number of seeds where random numbers occurred

   Set these parameters.

4. Click the OK button.

5. Click the Phylogenic Tree button to start analysis.

3.41 Creating Multiple Alignment Profiles (Amino Acid)

This function creates profiles of multiple alignment.   The multiple alignment between input sequences is calculated in advance and saved as a profile.   This allows high-speed alignment calculation between an unknown sequence and the profile.   The Clustal W algorithm, developed by J. Thompson and T. Gibson, is used for alignment calculation.

What is a profile?
A multiple alignment profile is pre-calculated data for the alignments between multiple input sequences that is saved for later use.

Why do I want to use a profile?
Calculating multiple alignments requires a long time.   For example, in the case of DNA, DNASIS MAX requires only ten minutes to calculate multiple alignments for 40 data items, but it may require two days for 200 data items.   This applies when the average BP length for the input sequences is about 1.5Kbp.   Longer sequences, such as a gene or a complete genome, require a longer time.
If you have many known sequences and want to calculate alignment between an unknown sequence and the known ones, you can save the time required to calculate alignment with the unknown sequence by creating a profile first.
Calculating a profile requires the same time as an ordinary calculation.   However, once a profile is created, DNASIS MAX can calculate alignment with the unknown sequence much faster (in about 10 seconds for the above example).

Disadvantages of using a profile
Using a profile provides fast calculation.   However, it results in degraded alignment precision.   The same data may produce different results when you use a profile and do not use a profile.   You should consider those characteristics when using a profile.

Procedure for Creating a Profile

Like any other mode of analysis, click the Analysis menu when creating a profile.   Here is a list of precautions:

1. Read a sequence you want to create into the Main window.

2. Click the Create Multiple Alignment Profile button and an Analysis dialog box. Then click the Parameter button.
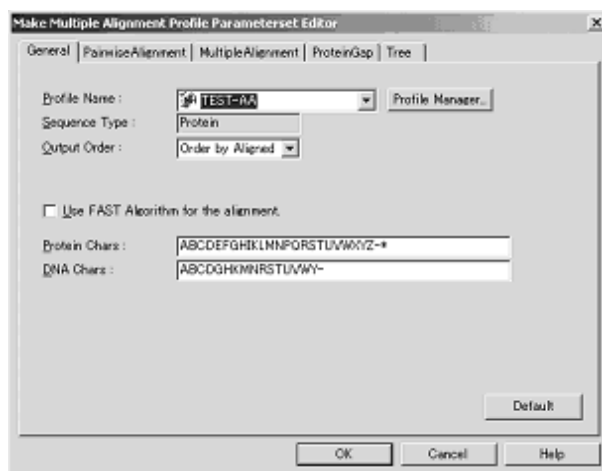
3. In the Profile Name field, enter a profile you want to create, and click the OK button.   To create a new profile, select Profile Manager... and use the Profile Manager*.

4. Click the Create Multiple Alignment Profile button.

DNASIS MAX uses all sequences displayed in the Sequence View to perform multiple alignment, and then writes the result into the profile.
Note: Because the profile is overwritten, be sure to make the profile setting before pressing the Analysis button.
Locking the profile prevents an unexpected overwrite.   Use the Profile Manager* for locking the profile.

Using a Created Profile on Another PC

You can export a newly created profile and save it to a file.   You can also import such an exported profile to use it on another PC.   The procedures are the same as those for DNA*.

3.42 Using Phylogenic Tree - Profiles (Amino Acid)

This function creates a phylogenic tree by adding a single sequence to a multiple alignment profile that has been produced in advance.

Analysis Procedure

1. Click the Phylogenetic Tree (using profile) icon from analysis button view and an Analysis dialog box will appear.

   Then click the Parameter button.
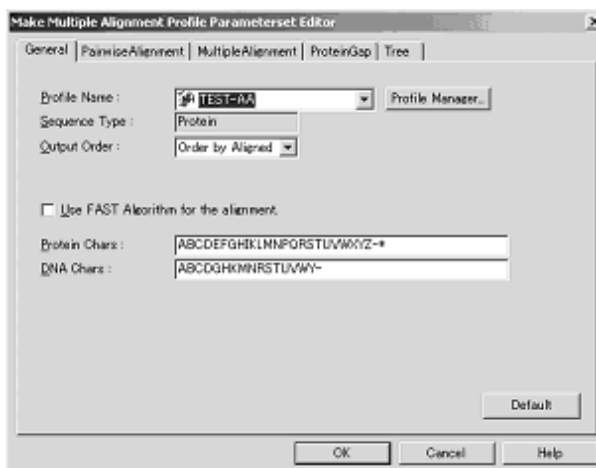
2. In the Profile Name field, enter a file you want to create, and then click the OK button.

3. In the Sequence View, select a sequence you want to analyze (a sequence to be added to a tree) as the target.

4. Click the Phylogenic Tree (using profile) button in the Analysis menu.

Explanation of the Result Window

When a branch is added to a phylogenic tree created using a profile, the branch is shown red.   For details about the window that displays phylogenic trees, refer to the Explanation of the Result Window in Section 23, "Phylogenic Tree-DNA".
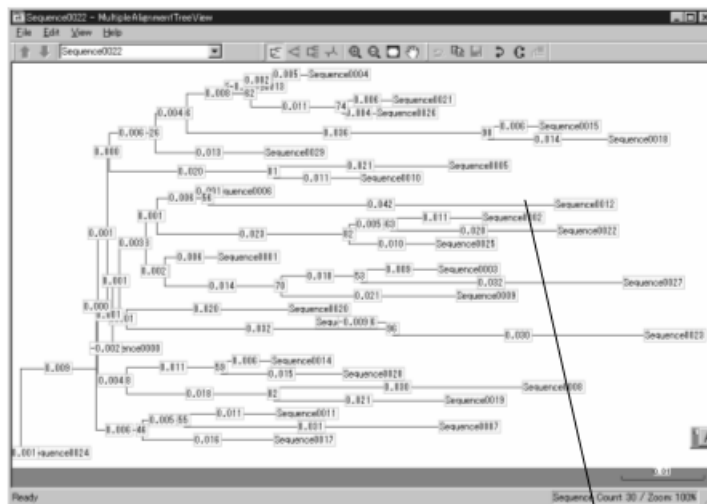
Added sequence

3.43 NCBI Entrez Search

This function connects the NCBI's Web site and performs entry search based on keywords from the Entrez database.   It also produces a list of accession numbers and definitions as the result of analysis.

Selecting a sequence record from this list results in importing the sequence into the Project.

Since this function directly connects to the NCBI's Web site, you need to set the Internet configuration parameters.

When using a proxy server, you also need to set HTTP Proxy in the Internet Options*.

Explanation of the Search Window

Item name (Parameter name)	Description
Database	Selects the type of the database as the target of search (Nucleotides/Proteins).
Operator	Chooses from OR, AND, and Delete this (deleting one line)
Search target field	Sets the field as the target of search.
Joining condition	Chooses from the following: is, is not, begin w/, and does not begin w/.
Value input field	Enters a search word, date, number, and other data.
New Keyword button	Adds one line to the keyword. You can set up to 20 lines.
Clear All button	Deletes all keywords that have been set.
Search button	Starts search under the preset search condition.
Options... button	Opens the dialog box that enables option settings.
Close button	Closes the dialog box without performing search. What has been entered is saved.
Cancel button	Closes the dialog box without performing search. What has been entered is not saved.
Help button	Displays online help.

Operators

Search by specifying multiple search conditions requires you to set logical operators that connect condition equations.

Operator	Format	Description
AND	<condition-equation-1> AND <condition-equation-2>	Searches for the entries meeting all condition equations connected by an AND operator.
OR	<condition-equation-1> OR <condition-equation-2>	Searches for the entries meeting either of the condition equations connected by an OR operator.

Join Conditions

Join conditions are used to set comparative operators between the items and values of condition equations.

Joining condition	Description
is	Searches for the entries having the same value as the setting in the Value Input field. If a word is entered, the entries that exactly match the whole word are hits; partial matches are ignored. Entering two or more words means that they are considered to form a phrase; therefore, the entire phrase is the candidate for a hit.
is not	Negates the meaning of the verb "is": that is, searches for the entries having a value not equal to

	the setting in the Value Input field.
begin /w	Searches for the entries having a word that begins with the character string specified in the Value Input field. Entering more than one word will result in improper search.
dose not begin /w	Negates the meaning of the verb "begin /w".

Option Setting Dialog Box
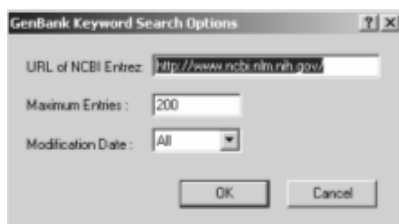
Item name (Parameter name)	Description
URL of NCBI Entrez	Set the URL of the NCBI site.
Maximum Entries	Sets the number of entries actually obtained from hits (not affecting the number of hits displayed).
Modification Date	Sets the dates of entries actually obtained from hits (not affecting the number of hits displayed).

Explanation of the Result Window

Item name (Parameter name)	Description
Result list	You can select the results by clicking them. You can use the Shift or the Ctrl key to select more than one entry. From the menu bar, you can also choose Edit and then Select All to select all entries. With an entry selected, choosing Edit and then Copy copies tab-delimited text data of the accession number and definition to the clipboard. You can paste this data to MS-Excel or other applications.
Column header	Clicking the column header with the mouse makes it possible to sort data based on the column. Another click on a key column toggles between ascending and descending order.
GenBank Report button	Obtains the layout of selected entries. The obtained data is added to the Sequence View.
New Search button	Deletes all the current search conditions and enters new search conditions.

Search Within Results button	Displays the search parameter input window for narrow-down search while keeping the current search conditions.

The search results are stored. The stored results can be retrieved from the next data list window.
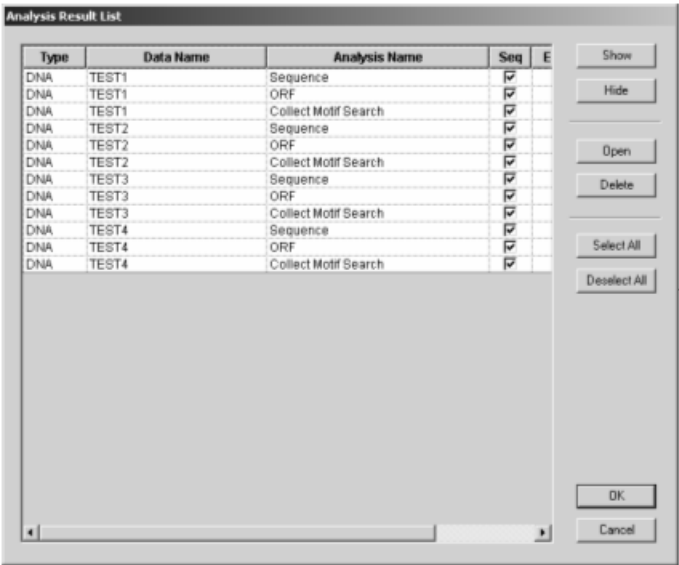
3.44 Searches Using GeneIndex

Use GeneIndex to perform Homology Search as well as Motif and Domain Search.   Performing a search requires an internet connection.   For operating environment details, refer to **GeneIndex 2.2 Operation Manual**.

Obtaining Accounts

The search engine that GeneIndex uses requires two accounts.   A DNASIS account that is included in the DNASIS MAX package and a GeneIndex contract account.
The DNASIS account and password are located inside the DNASIS MAX package.   If you do not find it inside the package please contact our support center.
Also if you are making a new GeneIndex contract, before using the function for the first time, you must first log in to the account from the website (http://index.dnasis.jp), agree to the contract terms and change the default password.
The number of users that can log in at the same time will depend on the contract options of your GeneIndex account. If you try to log in when number of users has already reached a maximum, an error message will appear.   When you leave the website be sure to log out.

Set GeneIndex Server Information

Before performing a search, you need to set login information for GeneIndex Server.   The login setting you make here is common to Homology Search and Motif and Domain Search.   After setting once the first time you will not have set up again each time you log in.
If you use a proxy to connect to the internet, you must set the proxy separately.   For details, refer to "1.6 Internet Settings Dialog Box".
Procedure:

1. Click the Option tab from analysis button view and click the GeneIndex Homology Search icon or GeneIndex Motif and Domain Search icon.
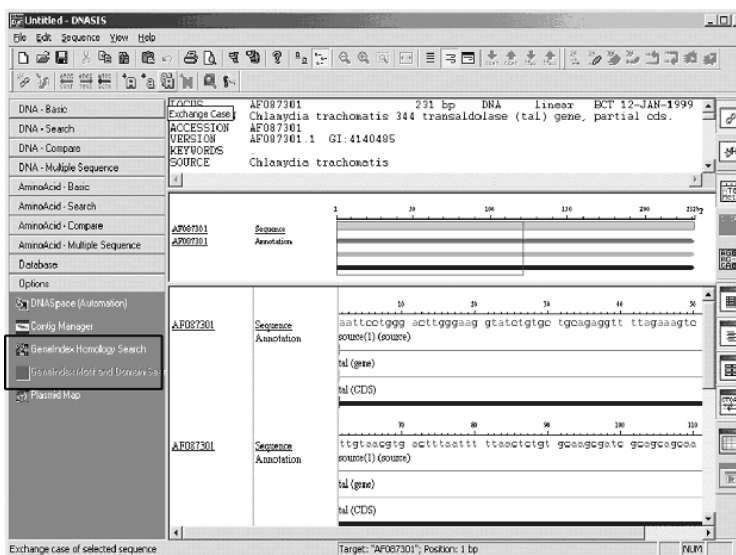
2. Click the Parameter... button from the Analysis dialog that appears.
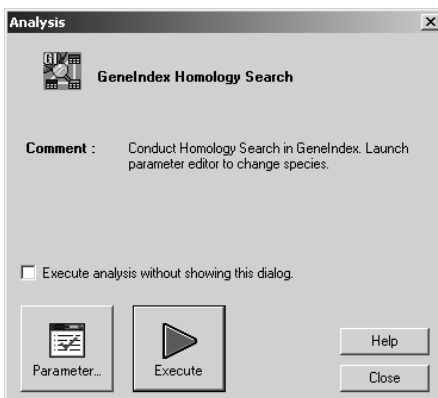
3. A settings dialog will pop up where you can enter GeneIndex login information.   Enter the appropriate information.
   (When you start up for the first time no species is registered in the species list combobox.   After entering the

appropriate value click the Update button and it will then be possible to access the GeneIndex server to obtain a species list.)
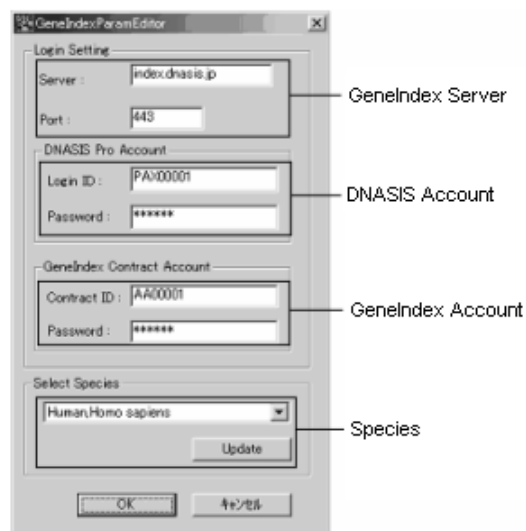
4. And click OK.

Homology Search

Use amino acid sequences and DNA sequences to perform a GeneIndex Homology Search.   When using this search, a sequence that displays in DNASIS MAX will appear in the search string of the GeneIndex Homology Search.   It is possible to specify other conditions then perform the search.

Select Target Sequences

The sequences that display in Sequence View are the target in the Homology Search.   You will have to hide sequences you want to remove from the search target.

1. Enter a new sequence or import one from an existing file.

2. It is possible to select either a DNA sequence or amino acid sequence.   For selecting a DNA sequence, click [icon] from the View Toolbar.   For selecting an amino acid, click [icon].

3. To remove a sequence from the search, right click over the sequence name.

4. Then select Hide from the popup menu.   The sequence will be hidden, and only the search target sequences will display in Sequence View.

Select Species

Select a target species for the Homology Search.

1. Click the Options item from the left-hand vertical menu.

2. Click the GeneIndex Homology Search icon then click the Parameter… button in the Analysis dialog box that appears. Select the target species and click OK.

Perform Homology Search

1. The target sequences for the search will display in Sequence View.

2. Click the Option tab from analysis button view and click the GeneIndex Homology Search icon.
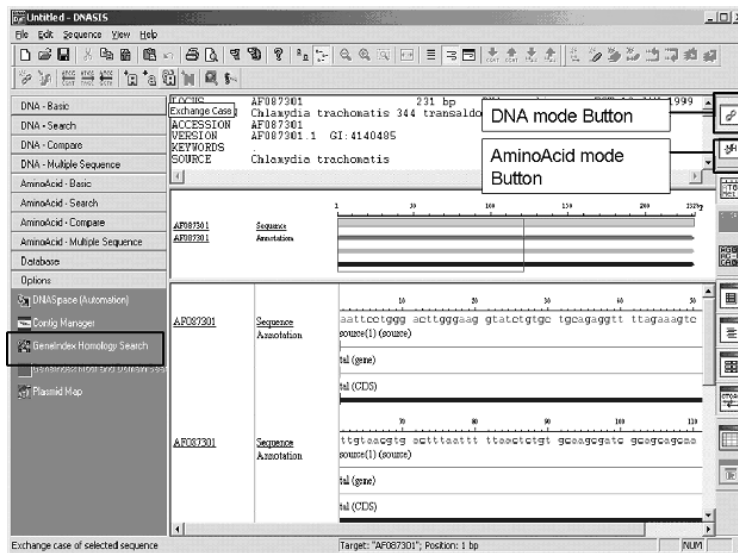
3. Click the Execute button from the Analysis dialog that appears.
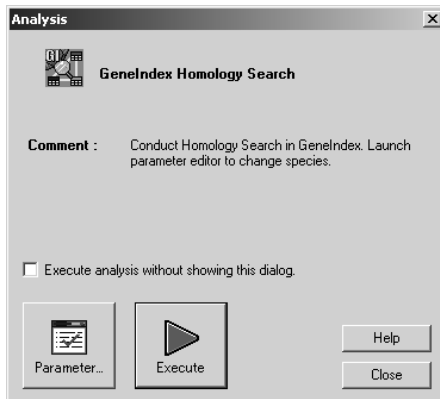
4. A browser will open and connect you to the GeneIndex site.   A database selection page will appear.   Select target databases for the GeneIndex Homology Search, and click Next.

5. A page for entering search conditions will appear.   For the sequence string, the sequence shown in DNASIS MAX will display here in FASTA format. Enter any other conditions and click Search.   The search results will appear.

6. An Export to DNASIS button will appear in the Search Result window so click it and a DNASIS Export window will appear.

7. Set the export parameter and click the Export to DNASIS button again to start the download.

Upper Limit of Characters

In Homology Search, the maximum number of searchable characters after converting to FASTA format is 20,000. Each sequence consists of ">" at the head, title, and sequence itself followed by a linefeed.   That is, five characters are automatically added to a sequence.   Even for multiple sequences, the total maximum length is 20,000.

For example, if displayed in Sequence View as below, it will be converted into FASTA format as below.   In this case, the number of characters is 38.

Sequence View

After conversion to FASTA format

```
>DNA001
aaggttcc
>DNA002
atgcatgc
```

If the upper limit error dialog appears in multiple sequence view, decrease the target sequences, refer to "Select Target Sequences" in "Homology Search" of "3.44 Searches Using GeneIndex", and try Homology Search again.

Motif and Domain Search

Use amino acid sequences to perform GeneIndex Motif and Domain Search.   When using this search, a sequence that displays in DNASIS MAX will appear in the search string of the Motif and Domain Search.   It is possible to specify other conditions then perform the search.

Select Target Sequences

The sequences that display in Sequence View are the target in the Motif and Domain Search.   You will have to hide sequences you want to remove from the GeneIndex search target.

1. Enter a new sequence or import one from an existing file.

2. Click ⊞ from the View Toolbar to switch to amino acid view.

3. To remove a sequence from the search, right click over the sequence name.

4. Then select Hide from the popup menu.   The sequence will be hidden, and only the search target sequences will

   display in Sequence View.

Select Species

Select a target species for the Motif and Domain Search.

1. Click the Options item from the left-hand vertical menu.

2. Then right click the GeneIndex Motif and Domain Search icon and select Parameter... Select the target species and

   click OK.

Perform Motif and Domain Search

1. The sequences you want to search will display in Sequence View.

2. Click the Option tab from analysis button view and click the GeneIndex Motif And Domain Search icon.
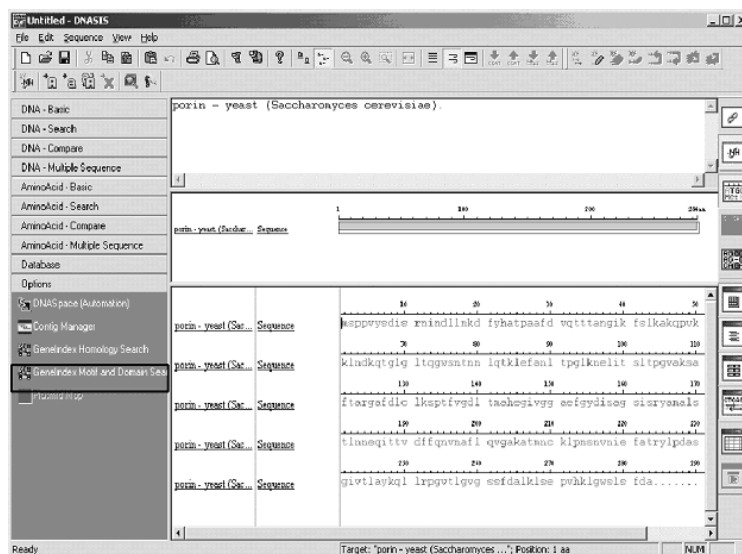
3. Click the Execute button from the Analysis dialog that appears.
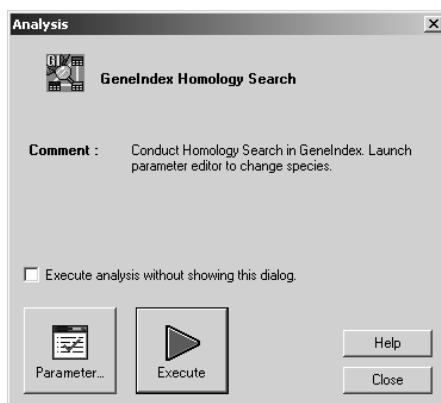
4. A browser will open and connect you to the GeneIndex site.   A database selection page will appear.   Select target databases for the GeneIndex Motif and Domain Search, and click Next.

5. For the sequence string, the sequence shown in DNASIS MAX will display here in FASTA format.   And click Search.   The search results will appear.

6. An Export to DNASIS button will appear in the Search Result window so click it and a DNASIS Export window will appear.

7. Set the export parameter and click the Export to DNASIS button again to start the download.

If multiple sequences were displayed in DNASIS MAX, they will also display in FASTA format under GeneIndex search conditions.   However, Motif and Domain Search will only return a result for the lead sequence.

Upper Limit of Characters
In Motif and Domain Search, the maximum number of searchable characters after converting to FASTA format is 20,000. For details, refer to "Upper Limit of Characters" in "Homology Search" of "3.44 Searches Using GeneIndex".

Export to DNASIS button
If you log in to GeneIndex from DNASIS MAX the Export to DNASIS button will appear in the Homology Search Result, Motif And Domain Search Result and Index Search Result windows.   If you click the Export to DNASIS button, it is possible to export an associated compressed file to DNASIS MAX from a homology search result or motif and domain search result.

Exporting to DNASIS MAX
The file downloaded when you click the Export to DNASIS button is compressed in LZH format.   Click Open from the dialog that normally appears when a download completes and an installer will start up DNASIS MAX and import the search result file.
If you click Save from the dialog that normally appears and a file with extension name dnasislzh will be saved to the folder you select. Below is the method to import a search result file into DNASIS MAX.
a. Double click on the search result file.
b. Drag and drop the search result file on the DNASIS MAX icon on your desktop.
c. Unzip the search result file with software that uses the UNLHA32.DLL then import the unzipped file.

Parameter Set List and Parameter Meanings
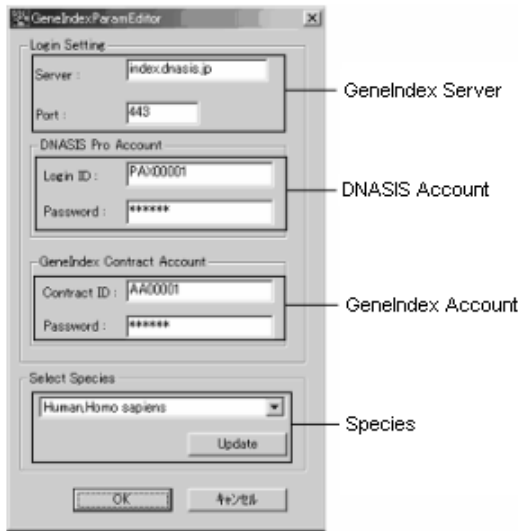
Item	Description
Server	Set GeneIndex Server name.
Port	Set the port number of GeneIndex Server.
DNASIS MAX Account : Login ID	Set the DNASIS account ID.
DNASIS MAX Account : Password	Set the DNASIS account password.
GeneIndex Contract Account : Contract ID	Set GeneIndex user name.
GeneIndex Contract Account : Password	Set GeneIndex password.
Select Species	Set the species. It is possible to select Human, Mouse and Rat. It is possible to select one from the list. The Update button renews the species list.

About GeneIndex 2.2

For details on how to use, refer to GeneIndex 2.2 Operation Manual.

Chapter 4 Details of Parameters

4.1 Complement Sequence

No parameters.

4.2 Reverse Complement Sequence

No parameters.

4.3 Reverse Sequence

No parameters.

4.4 Translation
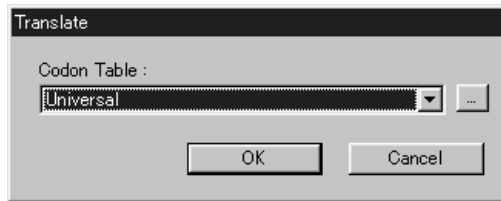
Item	Description
Codon Table	Selects a codon table used for translation.

4.5 Base Content

No parameters.

4.6 Codon Usage

Item	Description
Start Position	Specifies the base from which DNASIS starts counting codons. You can specify First, Second, or Third.

4.7 GC Content
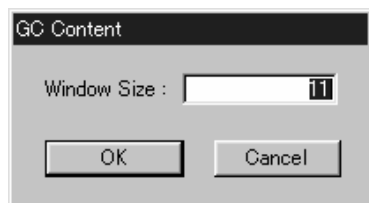
Item	Description
Window Size	Specifies the sequence window size for calculating the GC content.

4.8 Vector and Low-Quality End Trimming

Item	Description
Trim End	Specifies whether to trim the end. To trim the end, select the desired check box.
5' End	If this check box is selected, DNASIS trims the 5' end.
	If you select both "Trim at least..." and "Trim the first...", DNASIS will first trim as specified with "Trim at least..." and then trim as specified with "Trim the first...".
Trim at least...	Unconditionally trims the sequence of the specified length from the 5' end. Selecting the check box enables trimming. Specify an integer of 0 or greater as the sequence length.
Trim the first...	If this check box is selected, DNASIS trims the low-quality portion from the 5' end. Specify an integer of 0 or greater for the window length and quality threshold for determining quality.
	Trimming is performed as follows:
	1. Calculate the quality of the sequence (= window) of the specified length from the 5' end.
	2. If the quality is lower than the threshold in step 1, shift the window one base toward the 3' end and repeat step 1.
	3. When the quality becomes equal to or greater than the threshold in step 1, trim the portion starting from the 5' end and ending at the N that is closest to the 3' end within the current window.
3' End	If this check box is selected, DNASIS trims the 3' end. If you select both "Trim at least..." and "Trim the first...", DNASIS MAX will first trim as specified with "Trim at least..." and then trim as specified with "Trim the first...".
Trim at least...	Unconditionally trims the sequence of the specified length from the 3' end. Selecting the check box enables trimming. Specify an integer of 0 or greater as the sequence length.
Trim the first...	If this check box is selected, DNASIS MAX trims the low-quality portion from the 3' end. Specify an integer of 0 or greater for the window length and quality threshold for determining quality.
	Trimming is performed as follows:
	1. Calculate the quality of the sequence (= window) of the specified length from the 3' end.
	2. If the quality is lower than the threshold in step 1, shift the window one base toward the 5' end and repeat step 1.
	3. When the quality becomes equal to or greater than the threshold in step 1, trim the portion starting from the 3' end and ending at the N that is closest to the 3' end within the current window.
Same as 5' End	Specifies whether the conditions for trimming the 5' end are also applied to the 3' end. Selecting the check box causes DNASIS MAX to use the same conditions for the 5' and 3' ends.

Item	Description
Trim Vector	Specifies whether to trim a vector sequence. To trim a vector, select the check box.
Vector Name	Select the vector you want to trim from the list. You can select only one vector.
Cloning Site	Select the cloning site for the vector you want to trim from the list. You can select up to two items on the cloning site. To select more than one cloning site item, click each item in the list while holding down the Ctrl key.
	Specifies the minimum length of a match between the vector sequence and input sequence to use the DP method to determine the portion of the vector sequence to trim. If the matched length is smaller than this value, DNASIS MAX will not assume the sequence is a vector sequence and will not trim. Specify an integer of 15 or greater.
	Specifies the minimum match ratio between the vector sequence and input sequence to use the DP method to determine the portion of the vector sequence to trim. If the match ratio is smaller than this value, DNASIS MAX will not assume the sequence is a vector sequence and will not trim. Specify an integer of 0 or greater.
	Specifies the destination to output data if trimming results in a shorter sequence than the specified length.
	- If this check box is selected, DNASIS MAX will output data to the Others folder (lower folder).
	- If this check box is not selected, DNASIS MAX will output data to the Trimmed folder (upper folder). If "If vector trimming length is 0 bp..." is selected, however, that setting takes precedence.
	Specifies the destination to output data if vector trimming is not performed although it is specified.
Output Options	Not used in DNASIS MAX.
	Do not select the Output to "Others" folder check box. Selecting this check box will result in malfunction.
Default button	Returns parameters to default values.

4.9 ORF

Item	Description
Codon Table	Selects a codon table. To check the contents of a Codon Table, press "..." at the right edge of the Codon Table box. The Codon Table Editor will appear.
Initial Codons	Shows the initial codon name. The contents of the selected initial codon will be displayed. To check and edit the contents, click "...". The Initial Codon dialog will appear. The one with a check in the check box is the specified initial codon.

Open Reading Frame Search Result List
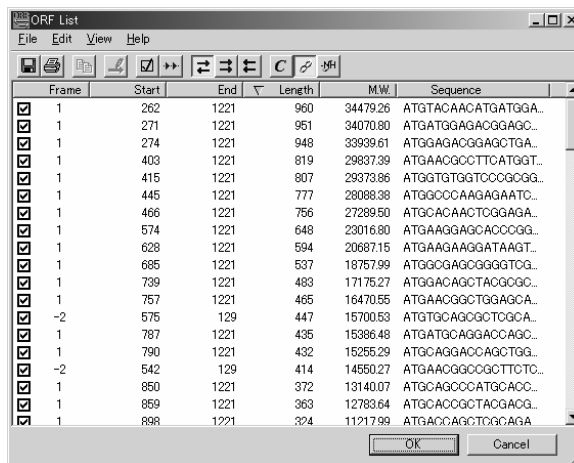
File menu	Description
Export…	Stores all ORF data (except for check boxes) as tab delimited text (with header) in a file.
Print Setup…	Sets the paper size to use for printing.
Print	Starts printing.
Close	Closes the search result list.

Edit menu	Description
Copy	Copies the selected ORF data (except for check boxes) to the clipboard as tab delimited text (with header).
Select All	Selects all the ORF data.
Check All	Check all the ORF data in the list.
Uncheck All	Uncheck all the ORF data ORF in the list.
Shape Setting…	Edits the status on Map View.

View menu	Description
Show Only Checked	Displays only the checked ORF data.
Show All Codons	Displays all the ORF data including the Start Codon Stop Codon not in reading frames.
All Frames	Displays all the frames.
Normal Frames	Displays the frames of normal strands.
Complementary Frame	Displays the frames of complementary strands.
Show Comment	Displays comments in the Result List.
Show DNA Sequence	Displays switching to DNA sequence.
Show Translated Sequence	Displays switching to amino acid translation sequence.

Help menu	Description
Help	Displays online help.

Toolbar

Icon	Description
	The same as selecting Export ... from File in the menu.
	The same as selecting Print Setup ... from File in the menu.
	The same as selecting Copy from Edit in the menu.
	The same as selecting Shape Setting ... from Edit in the menu.
	The same as selecting Show Only Checked from View in the menu.
	The same as selecting Show All Codons from View in the menu.
	The same as selecting All Frames from View in the menu.
	The same as selecting Normal Frames from View in the menu.
	The same as selecting Complementary Frame from View in the menu.
	The same as selecting Show Comment from View in the menu.
	The same as selecting Show DNA Sequence from View in the menu.
	The same as selecting Show Translation Sequence from View in the menu.

4.10 Primer Design

General Parameters

Item	Description
pick left primer	Specifies whether to design a left primer (5' upstream primer).
	To design the left primer, select the leftmost check box. You can also directly enter a sequence to specify the primer.
	You must specify a primer sequence in the direction of 5' to 3' on the input sequence.
pick right primer	Specifies whether to design a right primer (3' downstream primer).
	To design the right primer, select the leftmost check box. You can also directly enter a sequence to specify the primer. You must specify a primer sequence in the direction of 5' to 3' on the complement sequence of the input sequence.
pick hybridization probe	Specifies whether to design a hybridization probe for an amplification segment with a designed primer.
	To design the hybridization probe, select the leftmost check box. You can also directly enter a sequence to specify the probe. You must specify a probe sequence in the direction of 5' to 3' on the input sequence.
Sequence ID	Not supported by DNASIS MAX.
Target	Specifies a region or regions you want to have the PCR reaction product contain. Enter regions as follows:
	startbp,length startbp,length startbp,length...
	(Example 1)
	- Specifying "50,2" indicates that the product will contain 2bp from a position 50bp away from the 5' end, that is, bases 50-51bp.
	(Example 2)
	- Specifying "50,2 80,5" indicates that the product will contain 2bp from a position 50bp away from the 5' end and 5bp from a position 80bp away from the 5' end, that is, bases 50-51bp and 80-84bp. To specify more than one region, delimit regions with a space. When more than one region is specified, DNASIS MAX will design a primer which contains at least one of them. If you do not specify any region, DNASIS MAX will find the optimum primer from all regions of the input sequence.
Excluded Regions	Specifies a region or regions you want to exclude from the primer sequence search. Enter regions as follows:
	startbp,length startbp,length startbp,length...
	(Example 1)
	- Specifying "50,2" indicates that the product will contain 2bp from a position 50bp away from the 5' end, that is, bases 50-51bp.
	(Example 2)
	- Specifying "50,2 80,5" indicates that the product will contain 2bp from a position 50bp away from the 5' end and 5bp from a position 80bp away from the 5' end, that is, bases 50-51bp and 80-84bp.

Item	Description
	To specify more than one region, delimit regions with a space. When more than one region is specified, DNASIS MAX will design a primer and probe which do not overlap any of the regions.
Product Size	Specifies the minimum value (Min), optimum value (Opt), and maximum value (Max) for the length of the PCR reaction product.
Number To Return	Specifies the maximum number of primer candidates to be obtained.
Max Mispriming	Currently not supported.
Max 3' Stability	Specifies the maximum allowable value for Delta G necessary for duplex sequence dissociation at 5bp from the 3' end of the left primer and right primer. A larger value makes the 3' end more stable.
Pair Max Mispriming	Currently not supported.
Default button	Returns parameters to default values.

Primer Picking Conditions

Item	Description
Primer Size	Specifies the minimum value (Min), optimum value (Opt), and maximum value (Max) for the length of the primer sequence (bp). DNASIS MAX will not select primers shorter than the minimum value or longer than the maximum value. It will select a primer having the size closest to the optimum value. You cannot set the minimum value to 1 or less or the maximum value to greater than 36. (The maximum value of 36 is a limit due to the maximum sequence length when calculating the Tm value.) The minimum value cannot exceed the maximum value.
Primer Tm	Specifies the minimum value (Min), optimum value (Opt), and maximum value (Max) for the primer Tm value (Celsius). DNASIS MAX will not select primers having Tm lower than the minimum value or higher than the maximum value. It will select a primer having Tm closest to the optimum value.
Max Tm Difference	Specifies the maximum allowable value for the difference between Tm for the left primer and that for the right primer.
Product Tm	Specifies the minimum value (Min), optimum value (Opt), and maximum value (Max) for the Tm value (Celsius) for an amplification sequence (Product) with the designed primer. DNASIS MAX will not select a product having Tm lower than the minimum value or higher than the maximum value. When you specify the optimum value, DNASIS MAX will select the product having Tm closest to the optimum value if the Product Size for "Penalty for Primer Pairs" is other than 0. The product Tm value is calculated using the following formula: $Tm = 81.5 + 16.6(\log_{10}([Na^+])) + 0.41 \times (GC\%) - 600/length$ ([Na+]: sodium conc., GC%: GC content, length: sequence length) Primer GC% Specifies the minimum value (Min), optimum value (Opt), and maximum value (Max) for the primer GC content.
Max Self Complementarity	Specifies the maximum allowable value for an alignment score when local alignment is applied to a

Item	Description
	single primer or between the left and right primers.
	You can use this value to predict the trend in self-annealing for PCR.
	A score is calculated with the following values:
	EComplement base: +1.00
	- N: -0.25
	- Mismatch: -1.00
	- Gap: -2.00 (gaps larger than 2bp are now allowed)
Max 3' Self Complementarity	Specifies the maximum allowable value for an alignment score when 3' end alignment is applied to a single primer or between the left and right primers.
	You can use this value to predict the trend in primer dimer forming for PCR. A score is calculated in the same way as with Max Self Complementarity.
Max #N's	Specifies the maximum of number of Ns (undefined bases) that can be allowed for the designed primer.
Max Poly-X	Specifies the maximum number of consecutive identical bases (e.g., AAAAA).
Inside Target Penalty	Currently not supported.
Outside Target Penalty	Currently not supported.
First Base Index	Enter 1 for DNASIS.
GC Clamp	Designs a primer having a specified number of consecutive GCs at the 3' end of the left and right primers.
Salt Concentration	Specifies the salt (generally, KCl) concentration (mM) for PCR. Used to calculate a Tm value.
Annealing Oligo Concentration	Specifies the annealing oligo concentration (nM) for PCR. Used to calculate a Tm value.
Liberal Base	Selecting this check box enables DNASIS to accept a complex code, an asterisk (*), and a hyphen (-) by replacing them with Ns.

Pre-Sequence Inputs

Item	Description
Included Region	Specifies a region where you want to design a primer. Enter a region as follows:
	startbp,length
	Example: Specifying "50,451" indicates a 451bp region starting from 50bp, that is, 50-500bp. You cannot specify more than one region.
Start Codon Position	Currently not supported.
Sequence Quality	Enter a list of integers delimited with a space. When specifying this parameter, you must enter exactly one quality for each base.

Min Sequence Quality	Specifies the minimum value for the sequence quality within the sequence to be a primer.
Min End Sequence Quality	Specifies the minimum value for the sequence quality within 5bp at the 3' end of the primer.
Sequence Quality Range Min	Specifies the minimum value for the valid sequence quality.
Sequence Quality Range Max	Specifies the maximum value for the valid sequence quality.

Penalty Weights for Primer

Item	Description
Penalty for Primers	
Tm	Specifies penalties for a lower Tm value (Lt) and a higher Tm value (Gt) than the optimum Tm value for the designed primer.
Size	Specifies penalties for a smaller length (Lt) and a greater length (Gt) than the optimum length of the designed primer.
GC%	Specifies penalties for a smaller GC% (Lt) and a greater GC% (Gt) than the optimum GC% for the designed primer.
Self Complementarity	Specifies a penalty for a larger self complementarity than the optimum.
#N's	Specifies a penalty for a greater number of Ns than the optimum.
Mispriming	Currently not supported.
Sequence Quality	Specifies a penalty for a lower sequence quality than the optimum.
End Sequence Quality	Specifies a penalty for a lower quality than the optimum at the 3' end of the primer.
3' Self Complementarity	Specifies a penalty for a higher self complementarity than the optimum at the 3' end of the primer.
Position Penalty	Specifies a general penalty relating to the primer position.
End Stability	Specifies a penalty for a lower stability than the optimum at the 3' end of the primer.
Penalty for Primer Pairs	
Product Size	Specifies penalties for a smaller size (Lt) and a greater size (Gt) than the optimum product size.
Product Tm	Specifies penalties for a lower Tm value (Lt) and a higher Tm value (Gt) than the optimum Tm value (Celsius) for the product.
Tm Difference	Specifies a penalty for different Tm values between primers.
Any Complementarity	Specifies a penalty for a higher complementarity than the optimum between primers.
Hyb Oligo Penalty Weight	Specifies a weight used to calculate penalties for a primer pair and a probe.
3' Complementarity	Specifies a penalty for a higher 3' end complementarity than the optimum between primers.
Pair Mispriming	Currently not supported.
Primer Penalty Weight	Specifies a weight used to calculate penalties for a primer pair.

Hyb Oligo Conditions

Item	Description
Hyb Oligo Excluded Region	Specifies a region or regions you want to exclude from the probe design when designing a probe. Enter regions as follows: startbp,length startbp,length startbp,length... Example: Specifying "50,2 80,5" indicates a 2bp region starting from 50bp and a 5bp region starting from 80bp, that is, 50-51bp and 80-84bp. You can specify more than one region delimited with a space.
Hyb Oligo Size	Specifies the minimum value (Min), optimum value (Opt), and maximum value (Max) for the designed probe size (bp).
Hyb Oligo Tm	Specifies the minimum value (Min), optimum value (Opt), and maximum value (Max) for the Tm value (degrees Celsius) for the designed probe.
Hyb Oligo GC%	Specifies the minimum value (Min), optimum value (Opt), and maximum value (Max) for the GC% value for the designed probe.
Hyb Oligo Self Complementarity	Specifies the maximum value for probe self complementarity.
Hyb Oligo Max 3' Self Complementarity	Specifies the maximum value for self complementarity at the 3' end of the probe.
Max #N's	Specifies the maximum of number of Ns (undefined bases) that can be allowed for the probe.
Hyb Oligo Max Poly-X	Specifies the maximum number of consecutive identical bases (e.g., AAAAA) within the probe.
Hyb Oligo Min Sequence Quality	Specifies the minimum value for the sequence quality within the probe sequence.
Hyb Oligo Max Mishyb	Currently not supported.
Hyb Oligo Salt Concentration	Specifies salt concentration (mM) used to calculate the Tm value for the probe.
Hyb Oligo DNA Concentration	Specifies the annealing probe concentration (nM) used to calculate the Tm value.

Penalty Weights for Hyb Oligo

Item	Description
Hyb Oligo Tm	Specifies penalties for a lower Tm value (Lt) and a higher Tm value (Gt) than the optimum Tm value (degrees Celsius) for the probe.
Hyb Oligo Size	Specifies penalties for a smaller size (Lt) and a greater size (Gt) than the optimum probe size (bp).
Hyb Oligo GC%	Specifies penalties for a lower GC% value (Lt) and a higher GC% value (Gt) than the optimum GC% for the probe.
Hyb Oligo Self Complementarity	Specifies a penalty for a higher probe self complementarity than the optimum.
Hyb Oligo #N's	Specifies a penalty for a larger number of Ns (undefined bases) than the optimum within the probe.
Hyb Oligo Mispriming	Currently not supported.
Hyb Oligo Sequence Quality	Specifies a penalty for a lower probe sequence quality than the optimum.

4.11 Oligo Probe Design

The parameters for this analysis are the same as "4.10 Primer Design" describes.

4.12 Restriction Enzyme Site Search

Item	Description
Enzyme Name(NAME)	Shows the name of restriction enzymes. Using the restriction enzyme with a check in the leftmost check box, the cut region will be analyzed.
Recognition Sequence(SITE_N/SITE_C)	Indicates the sequence that a restriction enzyme recognizes in the direction from 5' to 3'. The "!" part indicates the place to cut. When the recognized sequence is not in a palindrome structure, the sequence that the Normal strand recognizes and the sequence that Complementary strand recognizes are displayed separated by "/".
Bases	Indicates the base number of the recognized sequence.
Kind Of Cut	Indicates the shape of the cut performed by the restriction enzyme.
5'-extended	Cuts the sequence so that the 5' end is longer than the 3' end. 5'-G AATTC-3' 3'-CTTAA G-5'
3'-extended	Cuts the sequence so that the 3' end is longer than the 5' end. 5'-TGCG CA-3' 3'-AC GCGT-5'
blunt-cut	Cuts the sequence so that the 3' and 5' ends have the same length. 5'-CCC GGG-3' 3'-ACG CGT-5'
not identified	Indicates that the position to cut cannot be identified for the restriction enzyme. Even when this enzyme is checked, it will not be registered as a parameter.

Button	Description
Show All	Displays all the restriction enzymes in the database.
Show Selected	Displays all the selected restriction enzymes.
Show Checked	Displays all the checked restriction enzymes.
Show Unchecked	Displays all the unchecked restriction enzymes.
Check All	Check all restriction enzymes.
Uncheck All	Uncheck all restriction enzymes
OK	Sets the checked restriction enzymes to the enzymes that the method will use, and exits from the Parameter Set Editor.
Cancel	Exits from the Parameter Set Editor without saving changes to the parameters.
Restriction Enzyme Database Manager	Starts the Restriction Enzyme Database Manager.

Restriction Site Param Editor dialog

Item	Description
Select from a Category	Check to select from a category.
Recognition Length	When selected, the restriction enzymes of recognition sequence lengths of 4, 5, 6, and 7 bp will be selected additionally.
Cut Kind	When selected, the restriction enzymes with a cut of 5'-extended, 3'-extended, and blunt-cut will be selected additionally.
Select From a List	Check to select restriction enzymes from the list.
User Selected	Displays the number of selected restriction enzymes.
...Button	Displays the Restriction Enzyme dialog (the previous diagram).
Number of Cutting Sites per enzyme	Limits the searched restriction enzymes within the upper and lower limits of the frequency of the cut places.
MIN	Specifies the lower limit of the frequency of the cut places.
MAX	When checked, the upper limit of the frequency of the cut places can be specified.
Help	Displays online help.
OK	Saves the set contents and exits the Restriction Site Param Editor dialog.
Cancel	Exits the Restriction Site Param Editor dialog without saving the set contents.

4.13 Motif Search
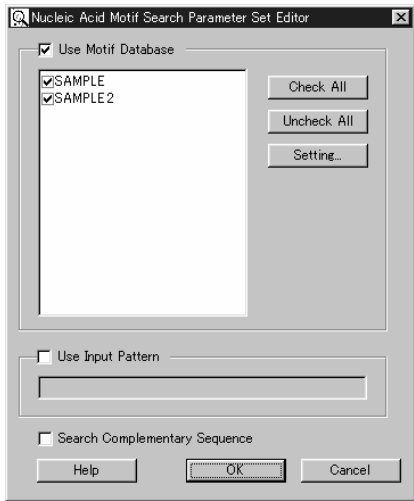
Item	Description
Use Motif Database	Select this check box when using a motif database. If this check box is selected, you must select at least one motif database. You must select either or both of the Use Input Pattern and Use Motif Database check boxes.
Motif Database list	A list of the motif databases registered in the database. Select the check boxes of the databases you want to search. If the list does not display any motif databases, click the Setting... button to start the Nucleic Acid Motif Search Database Manager, and specify the directory containing databases.
Use Input Pattern	Select this check box to search for a motif using an input pattern. If this check box is selected, you must enter a pattern. You must select either or both of the Use Input Pattern and Use Motif Database check boxes.
Input Pattern text box	Enter a search pattern when searching for a motif using an input pattern.
Search Complementary Sequence	Select this check box if you also want to search for a motif for a Complementary sequence. If this check box is cleared, DNASIS will only search for a motif for a Normal sequence. If this check box is selected, DNASIS MAX will search for a motif for both Normal and Complementary sequences.

Button	Description
Check All button	Selects all the checkboxes for the motif databases.
Uncheck All button	Unselects all the checkboxes for the motif databases.
Setting… button	Starts the Nucleic Acid Motif Search Database Manager for setting database directories.
Default button	Returns parameters to default values.

4.14 Mutational Site Search

Item	Description
Enzyme Name(NAME)	Name of the restriction enzyme. DNASIS analyzes the cut region using the restriction enzyme selected with the check box on the left.
Recognition Sequence (SITE_N/SITE_C)	Indicates the sequence that the restriction enzyme recognizes in the direction from 5' to 3'. The exclamation mark (!) indicates the position to cut. If the recognition sequence does not have a palindrome structure, the sequence recognized by a Normal sequence and that recognized by a Complementary sequence are separated with a slash (/).
Bases	Indicates the number of bases in the recognition sequence.
Kind Of Cut	Indicates the shape of the cut performed by the restriction enzyme.
5'-extended	Cuts the sequence so that the 5' end is longer than the 3' end.
3'-extended	Cuts the sequence so that the 3' end is longer than the 5' end. 5'-GAATTC-3' 3'-CTTAAG-5'
blunt-cut	Cuts the sequence so that the 3' and 5' ends have the same length. 5'-TGCGCA-3' 3'-ACGCGT-5'
not identified	Indicates that you cannot identify the position to cut for this restriction 5'-CCCGGG-3' 3'-ACGCGT-5' enzyme. If you check this enzyme, it will not be registered as a parameter.
Restriction Enzyme Database Manager button	Starts the Restriction Enzyme Database Manager.
HTML Mode	Select this check box for DNASIS.
Restriction Codon Table	Specifies a codon table to be used.

4.15 Hairpin Loop Search
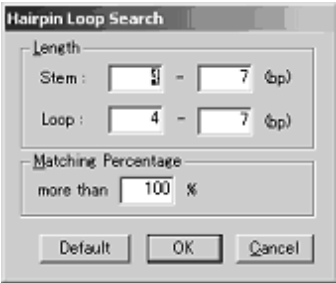
Item	Description
Stem Length	Specifies the stem length. If the stem length is within the range specified here, the stem length will become a hairpin loop region candidate.
	Input range: 2 to 99
Loop Length	Specifies the loop length. If the loop length is within the range specified here, the loop length will become a hairpin loop region candidate.
	Input range: 2 to 99
	Input range: 0 to 2,147,483,646
Matching Percentage	Specifies the match rate within a stem. If the match rate is above this specified rate, the stem will become a hairpin loop region candidate.
	Input range: 1 to 100
Default button	Returns parameters to default values.
OK button	Closes the dialog after the parameters have been set with the values entered in the dialog.
Cancel button	Closes the dialog without updating the parameters.

4.16 Stacking Site Search
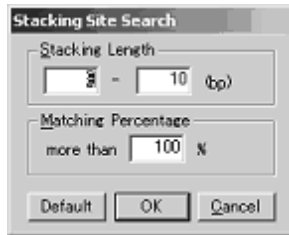
Item	Description
Stacking Length	Specifies the stacking site length. If the stacking site length is within the range specified here, the stacking site will become a hairpin loop region candidate. Input range: 2 to 99
Matching Percentage	Specifies the match rate within a stacking site. If the match rate is above this specified rate, the stacking site will become a hairpin loop region candidate. Input range: 1 to 100
Default button	Returns parameters to default values.
OK button	Closes the dialog after the parameters have been set with the values entered in the dialog.
Cancel button	Closes the dialog without updating the parameters.

4.17 Tandem Repeat Search

Item	Description
Repeat Length	Sets the repeat length. If the repeat length is within the range specified here, the repeat will become a hairpin loop region candidate. Input range: 2 to 99
Repeat Count	Specifies the number of repeat regions. If the number of repeat regions is above this number, the repeat will become a tandem repeat region candidate. Input range: 2 to 2,147,483,646
Default button	Returns parameters to default values.
OK button	Closes the dialog after the parameters have been set with the values entered in the dialog.
Cancel button	Closes the dialog without updating the parameters.

4.18 BLAST Search (DNA and Amino Acid)

Item	Description
Program name (PROGRAM)	Specifies the name of the program the BLAST search uses. Select one of blastp, blastn, blastx, tblastn, and tblastx. Use the following settings: DNA-BLAST search: blastn DNA-BLAST search (protein DB): blastx DNA-BLAST search (translation DB): tblastx Amino acid-BLAST search: blastp Amino acid-BLAST search (translation DB): tblastn
Detail	Displays the Detail dialog box.
Expectation value (E_VALUE)	Specifies an expectation value. DNASIS MAX will only report hits having an expectation value equal to or lower than the value specified here.
Filter query sequence (FILTER)	Specifies whether the same sequence as the input sequence will be excluded from the search target during database search. To exclude the same sequence, select this check box.
Include gap in alignment (INSERT_GAP)	Specifies whether to include a gap in alignment.
Descriptions(Descriptions)	DNASIS MAX to output up to a specified number of entries if the search results contain a large number of entries.
Alignments(Alignments)	Instructs DNASIS MAX to output up to a specified number of alignments if the search results contain a large number of alignments.
Nucleotide Database/Amino Acid Database (Target Databases)	Displays nucleic acid or amino acid databases depending on the type of the program. Select the check boxes of the databases you want to search.
Select All	Selects the check boxes of all databases.
Deselect All	Clears the check boxes of all databases.
Setting	Opens the BLAST Database Manager dialog box. You can change the directory in which DNASIS MAX will store the databases.
Default	Resets the parameters to their initial settings.

4.19 Internet BLAST Search (DNA and Amino Acid)
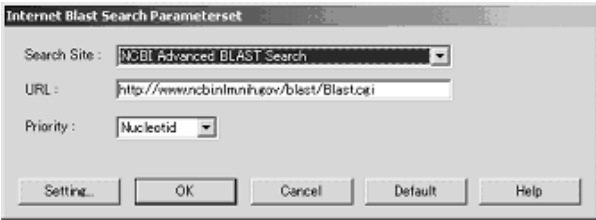
Item	Description
Search Site	Name of the search site. (This version only supports "NCBI advanced BLAST search".)
URL	URL of the NCBI site (used in case the URL is modified)
Priority	Specifies whether DNASIS will assume a sequence as a nucleic acid or amino acid sequence if it cannot determine the acid type from the sequence.
Default button	Returns parameters to default values.

Setting

In the Internet BLAST Search Parameterset dialog box, clicking the Setting button displays the following dialog box. You can specify search conditions when using the NCBI site to perform homology search.
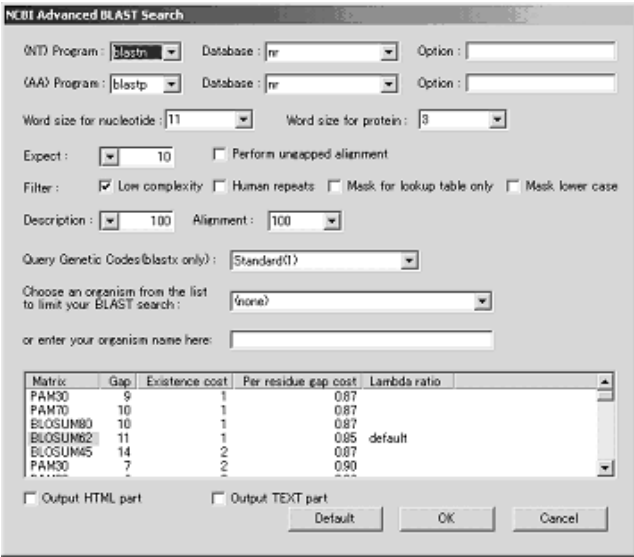
Note: For details about the parameters, refer to the NCBI Web site "Advanced BLAST".

4.20 Smith-Waterman Search (DNA and Amino Acid)

Item	Description
Calculate by	Specifies whether calculation based on the Smith-Waterman algorithm will be performed using the GENE BRIGHT III board or using software. GENE BRIGHT III: Use the board. You can select this item only when the board is available. Selecting this item displays the number of boards installed on the machine. Software Emulation: Performs calculation using software without using a board. Calculation by software will require very long time. Therefore, you should not select this item for a large database, for example, containing one million entries.
Input Sequence Type	Specifies the type of the input sequence. DNA: The input is a nucleic acid sequence. Amino Acid: The input is an amino acid sequence.
Matrix	Specifies the matrix used to calculate a score.
Initial Gap	Specifies a penalty score for inserting a gap. The value must be an integer between -16 and 0 inclusive and must not exceed the setting of Extension Gap.
Extension Gap	Specifies a penalty score for extending a gap. The value must be an integer between -16 and 0 inclusive and must not be smaller than the setting of Initial Gap.
Max Num of Return	Specifies the maximum number of results you want to obtain. Specify an integer between 1 and 500 inclusive.
Cut Off Score	If you select this check box, DNASIS will only output hits having a score larger than the specified score. If you do not select this check box, DNASIS MAX will output hits regardless of the score.
Target Database	Specifies the databases to be searched. The list box displays database name. Select the check boxes of the databases you want to search.
Default button	Returns parameters to default values.

4.21 Multiple Alignment (DNA and Amino Acid)

General

Parameter	Description
Sequence Type	"DNA" or "Protein" is automatically selected depending on the type of the selected profile. Note: Clustal W can process both DNA sequence data and amino acid sequence data. This parameter causes Clustal W to handle the input sequence data assuming it to be either a DNA or amino acid sequence. If this parameter is set to DNA, inputting amino acid sequence data (as determined by DNASIS MAX) results in an error and the sequence is not processed.
Output Order	Corresponds to the outorder parameter of Clustal W. You can select either 'Order by Aligned' or 'Order by Input', with which the value 'aligned' or 'input' will be set, respectively. This parameter specifies the order of the input sequence data within the result output from Clustal W. If you select 'Order by Aligned', DNASIS will output the results arranged in the order in which they appear in the guide tree or phylogenic tree. If you select 'Order by Input', DNASIS will output the results arranged in the order of the name of the input sequence data.
Use FAST Algorithm for the alignment	Corresponds to the quicktree parameter of Clustal W. Selecting the check box enables the parameter. Note: This parameter specifies whether or not to use a high-speed algorithm. If you select this parameter, DNASIS will use the Wilbur & Lipman algorithm to perform fast approximate processing. Without this parameter selected, DNASIS will use the Dynamic Programming algorithm to perform relatively slow but more accurate processing.
Protein Chars	Specifies the characters permitted within an input amino acid sequence . If the sequence contains any character other than those specified here, DNASIS does not recognize the sequence as an amino acid sequence. This parameter is used only for this application.
DNA Chars	Specifies the characters permitted within DNA sequence that are used to check input data. If the sequence contains any character other than those specified here, DNASIS MAX does not recognize the sequence as DNA sequence. This parameter is used only for this application.
Select create datas	Select Do Multiple Alignment only. For a phylogenic tree, select Do Tree only or Do Multiple Alignment and Tree.
Default button	Returns parameters to default values.

Pairwise Alignment

Item	Description
Gap Open Penalty	Corresponds to the pwgapopen parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting.
Gap Extension Penalty	Note: This parameter determines the probability of a gap (-) being inserted. A larger value makes it more difficult to insert a gap. Corresponds to the pwgapext parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter determines the probability of a gap being extended. A larger value results in a shorter gap.
Protein Weight Matrix	Corresponds to the pwmatrix parameter of Clustal W. You can select 'BLOSUM 30', 'PAM 350', 'Gonnet 250', or 'Identity matrix', which sets the value 'blosum', 'pam', 'gonnet', or 'id', respectively. Alternatively, you can select 'User defined' to have, DNASIS use the matrix file specified in the edit box. Note: This parameter specifies a table indicating similarity among amino acid molecules.
DNA Weight Matrix	Corresponds to the pwdnamatrix parameter of Clustal W. You can select 'IUB' or 'CLUSTALW(1.6)', which sets the value 'iub' or 'clustalw', respectively. Alternatively, you can select 'User defined' to have DNASIS use the matrix file specified in the edit box. Note: This parameter specifies a table indicating scores which specify whether DNA matches or does not match.
Gap Penalty	Corresponds to the pairgap parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: When using a high-speed algorithm, use this parameter to specify the Open and Extension gaps. The setting will not affect the processing speed unless you specify an extreme value.
K-tuple(word)size	Corresponds to the ktuple parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter specifies the size of a completely matched sequence. A larger value results in faster calculation. A smaller value results in higher precision.
No. of top diagonals	Corresponds to the topdiags parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: Clustal W calculates the number of complete matches within each diagonal (matched position in the sequence) and uses the matches having high match ratios for alignment. This parameter determines the number (n) of completely matched positions to be used; the n highest match ratios will be used. A smaller value results in higher precision. A larger value results in higher speed.
Window size	Corresponds to the window parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter specifies the number of diagonals around the completely matched portion that are used for alignment. A smaller value results in higher precision. A larger value results in higher speed.

Multiple Alignment

Item	Description
Gap Open Penalty	Corresponds to the gapopen parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter determines the probability of a gap being inserted. A larger value makes it more difficult to insert a gap.
Gap Extension Penalty	Corresponds to the gapext parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter determines the probability of a gap being extended. A larger value results in a shorter gap.
Delay divergent sequences	Corresponds to the maxdiv parameter of Clustal W. Note: This parameter prevents DNASIS from aligning sequences having distant relationships until it aligns the sequences having the closest relationship.
DNA Transitions Weight	Corresponds to the transweight parameter of Clustal W. Note: This parameter specifies a value of 0 or 1 for replacement. If 0 is specified, DNASIS does not assume replacement as a match. If 1 is specified, DNASIS assumes replacement as a match. You should specify 0 for closely-related DNA sequence data and 1 for distantly-related DNA sequence data.
Protein Weight Matrix	Corresponds to the matrix parameter of Clustal W. You can select 'BLOSUM series', 'PAM series', 'Gonnet series', or 'Identity matrix', which sets the value 'blosum', 'pam', 'gonnet', or 'id', respectively. Alternatively, you can select 'User defined' to have DNASIS use the matrix file specified in the edit box. Note: This parameter specifies a table indicating similarity among amino acid molecules.
DNA Weight Matrix	Corresponds to the dnamatrix parameter of Clustal W. You can select 'IUB' or 'CLUSTALW(1.6)' which sets the value 'iub' or 'clustalw', respectively. Alternatively, you can select 'User defined' to have DNASIS use the matrix file specified in the edit box. Note: This parameter specifies a table indicating scores which specify whether DNA matches or does not match.
Use negative matrix	Corresponds to the negative parameter of Clustal W. Selecting the check box enables the parameter. Note: Initially, a positive matrix is used. If this parameter is selected, a negative matrix is used.

Protein Gap

Item	Description
Residue-Specific gap off	Corresponds to the nopgap parameter of Clustal W. Selecting the check box enables the parameter. Note: Specify GapPenalty for each amino acid. A gap is likely to be inserted where many amino acids are set in the sequence data.
Hydrophilic gap off	Corresponds to the nohgap parameter of Clustal W. Selecting the check box enables the parameter. Note: Specifying this parameter increases the probability that a gap is inserted if five or more hydrophilic amino acids are contained consecutively.
Hydrophilic Residues	Corresponds to the hgapresidues parameter of Clustal W. Note: Specifying this parameter reduces the probability that a gap is inserted if gaps are too close to each other. A penalty is given if gaps are closer to each other than the value specified here.
Gap Separation Distance	Corresponds to the gapdist parameter of Clustal W.
End Gap Separation	Corresponds to the endgaps parameter of Clustal W. Selecting the check box enables the parameter. Note: Specifying this parameter prevents a gap from being created at the end. This parameter is useful for a sequence that is estimated as not important biologically.

Tree

Item	Description
Phylogenetic tree	Corresponds to the tree parameter of Clustal W.
Bootstrap tree	Select this check box when evaluating the reliability of the tree using the bootstrap method. Corresponds to the bootstrap(=n) parameter of Clustal W.

Number of bootstrap	Corresponds to the bootstrap(=n) parameter of Clustal W.
Seed no.	Corresponds to the seed parameter of Clustal W.
Exclude positions with gaps	Corresponds to the tossgaps parameter of Clustal W.
Correct for multiple substitutions	Corresponds to the kimura parameter of Clustal W.

4.22 Phylogenic Tree (DNA and Amino Acid)

The parameters are the same as "4.21 Multiple Alignment" describes.

4.23 Creating Multiple Alignment Profiles (DNA and Amino Acid)

General

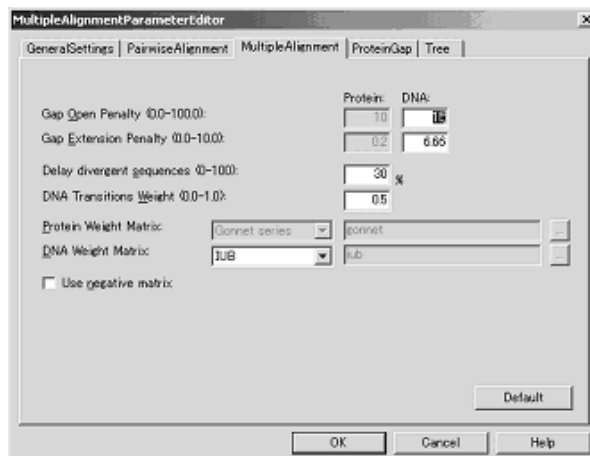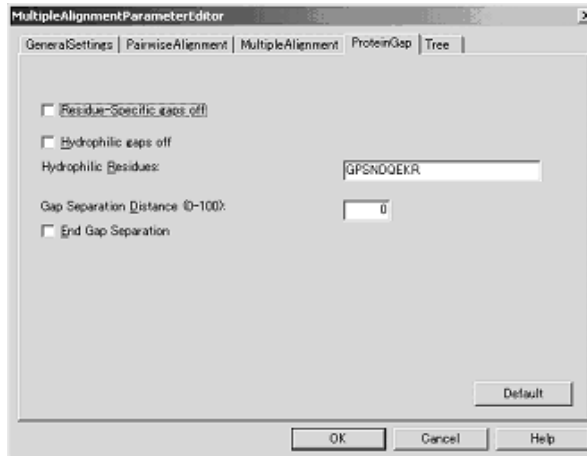| Item	Description
Profile Name	Specifies a profile for storing the results of multiple alignment calculation. To create a new profile, click the Profile Manager... button to open the Multiple Alignment Profile Manager. If there is another profile stored under the selected profile name, that profile will be overwritten. You cannot select a read-only profile. The [ ] icon represents a DNA sequence profile while the [-] icon represents an amino acid sequence profile.
Profile Manager...	Opens the Multiple Alignment Profile Manager. You can create, delete, import, and export a profile or modify the attributes of a profile.
Sequence Type	"DNA" or "Protein" is automatically selected depending on the type of the selected profile. Note: Clustal W can process both DNA sequence data and amino acid sequence data. This parameter causes Clustal W to handle the input sequence data assuming it to be either a DNA or amino acid sequence. If this parameter is set to DNA, inputting amino acid sequence data (as determined by DNASIS MAX) results in an error and the sequence is not processed.
Output Order	Corresponds to the outorder parameter of Clustal W. You can select either 'Order by Aligned' or 'Order by Input', with which the value 'aligned' or 'input' will be set, respectively. Note: This parameter specifies the order of the input sequence data within the result output from Clustal W. If you select 'Order by Aligned', DNASIS MAX will output the results arranged in the order in which they appear in the guide tree or phylogenic tree. If you select 'Order by Input', DNASIS MAX will output the results arranged in the order of the name of the input sequence data.
Use FAST Algorithm for the alignment	Corresponds to the quicktree parameter of Clustal W. Selecting the check box enables the parameter. Note: This parameter specifies whether or not to use a high-speed algorithm. If you select this parameter, DNASIS will use the Wilbur & Lipman algorithm to perform fast approximate processing. Without this parameter selected, DNASIS will use the Dynamic Programming algorithm to perform relatively slow but accurate processing.
Protein Chars	Specifies the characters permitted within an amino acid sequence that are used to check input data. If the sequence contains any character other than those specified here, DNASIS does not recognize the sequence as an amino acid sequence. This parameter is used only for this application.
DNA Chars	Specifies the characters permitted within DNA sequence that are used to check input data. If the sequence contains any character other than those specified here, DNASIS does not recognize the sequence as DNA sequence. This parameter is used only for this application.
Default button	Returns parameters to default values.

Pairwise Alignment

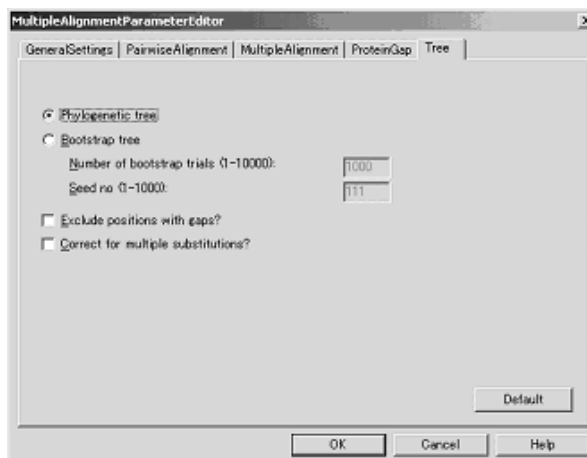| Item	Description
Gap Open Penalty	Corresponds to the pwgapopen parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter determines the probability of a gap (-) being inserted. A larger value makes it more difficult to insert a gap.
Gap Extension Penalty	Corresponds to the pwgapext parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter determines the probability of a gap being extended. A larger value results in a shorter gap.
Protein Weight Matrix	Corresponds to the pwmatrix parameter of Clustal W. You can select 'BLOSUM 30', 'PAM 350', 'Gonnet 250', or 'Identity matrix', which sets the value 'blosum', 'pam', 'gonnet', or 'id', respectively. Alternatively, you can select 'User defined', to have DNASIS use the matrix file specified in the edit box. Note: This parameter specifies a table indicating similarity among amino acid molecules.
DNA Weight Matrix	Corresponds to the pwdnamatrix parameter of Clustal W. You can select 'IUB' or 'CLUSTALW(1.6)' which sets the value 'iub' or 'clustalw', respectively. Alternatively, you can select 'User defined', to have DNASIS use the matrix file specified in the edit box. Note: This parameter specifies a table indicating scores which specify whether DNA matches or does not match.
Gap Penalty	Corresponds to the pairgap parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: When using a high-speed algorithm, use this parameter to specify the Open and Extension gaps. The setting will not affect the processing speed unless you specify an extreme value.
K-tuple(word)size	Corresponds to the ktuple parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter specifies the size of a completely matched sequence. A larger value results in faster calculation. A smaller value results in higher precision.
No. of top diagonals	Corresponds to the topdiags parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: Clustal W calculates the number of complete matches within each diagonal (matched position in the sequence) and uses the matches having high match ratios for alignment. This parameter determines the number (n) of completely matched positions to be used; the n highest match ratios will be used. A smaller value results in higher precision. A larger value results in higher speed.
Window size	Corresponds to the window parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter specifies the number of diagonals around the completely matched portion that are used for alignment. A smaller value results in higher precision. A larger value results in higher speed.

Multiple Alignment

Item	Description
Gap Open Penalty	Corresponds to the gapopen parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter determines the probability of a gap being inserted. A larger value makes it more difficult to insert a gap.
Gap Extension Penalty	Corresponds to the gapext parameter of Clustal W. Either the value for amino acid or that for nucleic acid is used depending on the Sequence Type setting. Note: This parameter determines the probability of a gap being extended. A larger value results in a shorter gap.
Delay divergent sequences	Corresponds to the maxdiv parameter of Clustal W. Note: This parameter prevents DNASIS from aligning sequences having distant relationships until it aligns the sequences having the closest relationship.
DNA Transitions Weight	Corresponds to the transweight parameter of Clustal W. Note: This parameter specifies a value of 0 or 1 for replacement. If 0 is specified, DNASIS does not assume replacement as a match. If 1 is specified, DNASIS assumes replacement as a match. You should specify 0 for closely-related DNA sequence data and 1 for distantly-related DNA sequence data.
Protein Weight Matrix	Corresponds to the matrix parameter of Clustal W. You can select 'BLOSUM series', 'PAM series' or 'Gonnet series', which sets the value 'blosum', 'pam', 'gonnet', or 'id', respectively. Alternatively, you can select 'User defined' to have DNASIS use the matrix file specified in the edit box. Note: This parameter specifies a table indicating similarity among amino acid molecules.
DNA Weight Matrix	Corresponds to the dnamatrix parameter of Clustal W. You can select 'IUB' or 'CLUSTALW(1.6)', which sets the value 'iub' or 'clustalw', respectively. Alternatively, you can select 'User defined' to have DNASIS use the matrix file specified in the edit box. Note: This parameter specifies a table indicating scores which specify whether DNA matches or does not match.
Use negative matrix	Corresponds to the negative parameter of Clustal W. Selecting the check box enables the parameter. Note: Initially, a positive matrix is used. If this parameter is selected, a negative matrix is used.

Protein Gap

Item	Description
Residue-Specific gap off	Corresponds to the nopgap parameter of Clustal W. Selecting the check box enables the parameter. Note: Specify GapPenalty for each amino acid. A gap is likely to be inserted where many amino acids are set in the sequence data.
Hydrophilic gap off	Corresponds to the nohgap parameter of Clustal W. Selecting the check box enables the parameter. Note: Specifying this parameter increases the probability that a gap is inserted if five or more hydrophilic amino acids are contained consecutively.
Hydrophilic Residues	Corresponds to the hgapresidues parameter of Clustal W. Note: Specifying this parameter reduces the probability that a gap is inserted if gaps are too close to each other. A penalty is given if gaps are closer to each other than the value specified here.
Gap Separation Distance	Corresponds to the gapdist parameter of Clustal W.
End Gap Separation	Corresponds to the endgaps parameter of Clustal W. Selecting the check box enables the parameter. Note: Specifying this parameter prevents a gap from being created at the end. This parameter is useful for a sequence that is estimated as not important biologically.

Tree

Item	Description
Phylogenetic tree	Corresponds to the tree parameter of Clustal W.
Bootstrap tree	Select this check box when evaluating the reliability of the tree using the bootstrap method. Corresponds to the bootstrap(=n) parameter of Clustal W.

Number of bootstrap	Corresponds to the bootstrap(=n) parameter of Clustal W.
Seed no	Corresponds to the seed parameter of Clustal W.
Exclude positions with gaps?	Corresponds to the tossgaps parameter of Clustal W.
Correct for multiple substitutions?	Corresponds to the kimura parameter of Clustal W.

4.24 Phylogenic Tree (Using Profiles (DNA and Amino Acid))

The parameters are the same as "4.23 Creating Multiple Alignment Profiles" describes.

4.25 Sequence Assemble

Item	Description
Parameterset Name	Display the parameterset name.
Parameterset Type	Display the type of parameterset.
Parameters	Display the parameters.
Mir_Overlap_Length	Minimum overlap length to assemble sequences. Overlap areas with shorter length than it are ignored. Range: 1 - 1000bp (The length shorter than "Homology_Compare_NA" is not allowed.)
Min_Match_Rate	Minimum matching rate to assemble sequences. Sequences with lower rate than it are not assembled. Range: 1 - 100%
Homology_Compare_NA	Minimum BPs compared in homology search. Homology search is conducted from/to the site with perfect match in the longer bases than it. Range: 1 - 6bp
MaxMatch_Compare_NA	Maximum BPs compared. If the length of compared bases is longer than it, the length is divided by 2 and compared again. Range: 200 - 500bp
Contig_Header	Contig Name Header. Contig name is Contig Name Header + making number. Character Range: 1 - 59 Valid Characters: Alphabet (Upper and Lower case) and "_".

4.26 Clustering

Item	Description
Sequence Type	Specifies the type of the input sequence: Nucleotides or Amino Acid.
Mode	Selects the type of clustering. The following two modes are available:
	Clustering only input sequences each other: Performs clustering between sequences in the Sequence Editor. An existing cluster-representing sequence database is not used.
	Clustering with existing cluster DB: Performs clustering between a sequence in the Sequence Editor and an existing cluster-representing sequence database. The cluster-representing sequence database will be updated as a result of clustering.
	If this mode is selected, the Browse button is enabled. Clicking the Browse button displays the database selection dialog box, in which you can select a cluster-representing sequence database.
Score is more than	Specifies a BLAST search score used as a basis for clustering.
Overlapping length for query length is more than	Specifies the ratio of the length of the matched sequence in BLAST search according to the entire query length of the sequence. This value is also used as a basis for clustering.

4.27 BLAST Search and Extraction
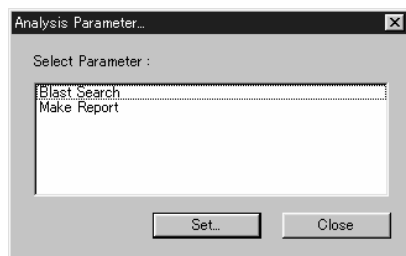
Item	Description
BLAST Search	Displays a dialog box used to set parameters for BLAST search.
Make Report	Displays a dialog box used to set parameters for extraction.

BLAST Search
Refer to "4.18 BLAST Search".

Make Report

Item	Description
Pick up top - items	Sorts the target list in descending order by Primary Keyword and Secondary Keyword (ascending order for E-Value) and extracts a specified number of entries from the top.
Primary Keyword:	Specifies a primary key used to sort the target list.
matching percentage:	Match ratio between the query sequence and target sequence at a homology matched portion.
Overlapping:	Number of bases matched between the query sequence and target sequence.
Score:	Score for a homology search.
E_value:	Expectation value for a homology search.
Secondary Keyword	Specifies a secondary key used to sort the target list. You can specify the same items as those for Primary Keyword.
Pick up all items with the same value	If this check box is selected, DNASIS MAX will extract all items that have both the value specified with Primary Keyword and that specified with Secondary Keyword even if the specified number of extracted items is exceeded.
Pick up items that	Extracts all items that satisfy the specified conditions from the results of homology search. You can specify conditions for each of the matching percentage, number of bases matched, score, and expectation value. All the conditions specified here are ANDed. If two or more conditions are specified, DNASIS MAX will extract the items that satisfy all those conditions.
matching percentage is more than	If this check box is selected, DNASIS MAX will extract targets having a matching percentage equal to or greater than specified value.

Item	Description
overlapping sequences are more than	If this check box is selected, DNASIS MAX will extract targets having a number of base matched equal to or greater than the specified value.
score is more than	If this check box is selected, DNASIS MAX will extract targets having a score equal to or greater than the specified value.
E_value is less than	If this check box is selected, DNASIS MAX will extract targets having an expectation value, equal to or greater than the specified value.
overlapping length for query length is more than	If this check box is selected, DNASIS MAX will extract targets having an "overlapping length for query length" equal to or greater than the specified value. The "overlapping length for query length" is the ratio of the number of matched bases to the query sequence length.
output header line	If this check box is selected, DNASIS MAX will add a header line to the output file.
output query sequence	If this check box is selected, DNASIS MAX will add a query sequence to the output file.

4.28 Amino Acid Content

No parameters.

4.29 Isoeletric Point

No parameters.

4.30 Hydrophilicity, Hydrophobicity, and Secondary Structure
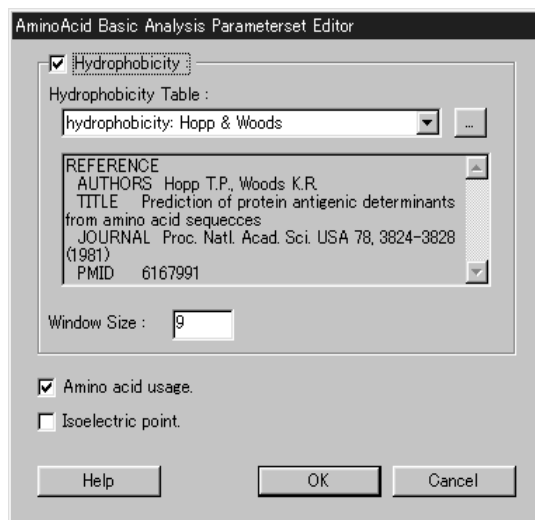
Item	Description
Hydrophobicity	Always select this check box.
Hydrophobicity Table	Specifies a table which defines hydrophilicity and hydrophobicity for each amino acid, used for hydrophilicity and hydrophobicity analysis.
Window Size	Specifies the window size when displaying hydrophilicity, hydrophobicity, and secondary structure.
Amino acid usage.	Do not select this check box.
Isoelectric point	Do not select this check box.

4.31 Proteolytic Site Search

Item	Description
Proteolytic Enzyme Name(NAME)	Name of the proteolytic enzyme. DNASIS MAX will perform analysis using the enzymes for which you select the check boxes on the left.
Recognition Sequence(SITE)	Sequence recognized by the proteolytic enzyme. An amino acid sequence is represented in the single-character format with an exclamation mark (!) indicating a cut position. If there are two or more recognition sequences, a slash (/) is used as a delimiter. If there are two or more recognition amino acids (complex code), they are enclosed by [ ]. X indicates any amino acid. Example: [KR] ! X / AR ! X Identify KX, RX, and ARX and cut between K and X, R and X, and AR and X.
Comment	Displays comments for the proteolytic enzyme, if any.
Check All button	Selects the check boxes of all the displayed enzymes.
Uncheck All button	Clears the check boxes of all the displayed enzymes.
Check Selected button	Selects the check boxes of all the selected proteolytic enzymes.
Uncheck Selected button	Clears the check boxes of all the selected proteolytic enzymes.
Show All button	Displays all proteolytic enzymes in the database.
Show Selected button	Displays all the selected the proteolytic enzymes.
Show Checked button	Displays all checked the proteolytic enzymes.
Show Unchecked button	Displays all unchecked the proteolytic enzymes.
Help button	Displays online help.
OK button	Sets the checked proteolytic enzymes to the enzymes that the method will use, and exits from the Parameter Set Editor.
Cancel button	Exits from the Parameter Set Editor without saving changes to the parameters.
Proteolytic Enzyme Database Manager button	Starts the Proteolytic Enzyme Database Manager.

4.32 Annotation

Annotation Setting dialog

Item	Description
Annotation Name	Shows the annotation name.
Annotation Kind	Shows the annotation type.
Link URL	Shows the annotation URL link.
Show Link...	Displays the webpage of the Link URL.
Annotation Range	Shows the annotation range.
Orient	Selects the annotation orientation.
Part Range	Shows the annotation part range.
Start	Shows the annotation start position.
End	Shows the annotation end position.
Add...	Displays the Add Annotation Part dialog to specify the part range.
Delete	Deletes the selected Part.
Edit...	Edits the selected Part.
Comment	Shows annotation comments.
Key	Shows comment keys.
Value	Shows comment values.
Add...	Displays the Add Annotation Comment dialog. Adds comments in the Add Annotation Comment dialog.
Delete	Deletes the selected comment.
Edit...	Edits the selected comment.
Line Width	Shows the line width.
Part Width	Shows the line width (horizontal) of Part.
Color	Shows the annotation color.
Color Setting...	Sets the annotation color.
OK	Sets the selected contents as parameters, and exits from the Annotation Setting dialog.
Cancel	Exits from the Annotation Setting dialog without saving changes to the parameters.

Kind Color Setting dialog
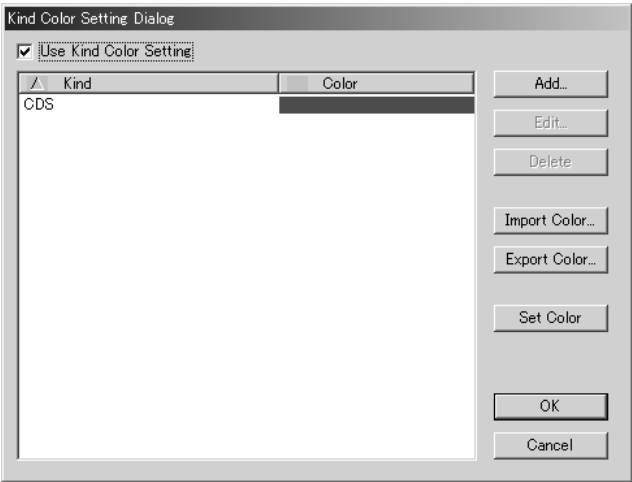
Item	Description
Use Kind Color Setting	Select to use the specified color setting.
Kind	Shows the type of color settings under Kind.
Color	Displays the type of color.
Add...	Adds color settings.
Edit...	Edits the selected color setting.
Delete	Deletes the selected color settings.
Import Color...	Imports color settings.
Export Color...	Exports color settings.
Set Color	Sets colors.
OK	Saves color settings and exits from the Kind Color Setting dialog.
Cancel	Exits from the Kind Color Setting dialog box without saving color settings.

Annotation List dialog
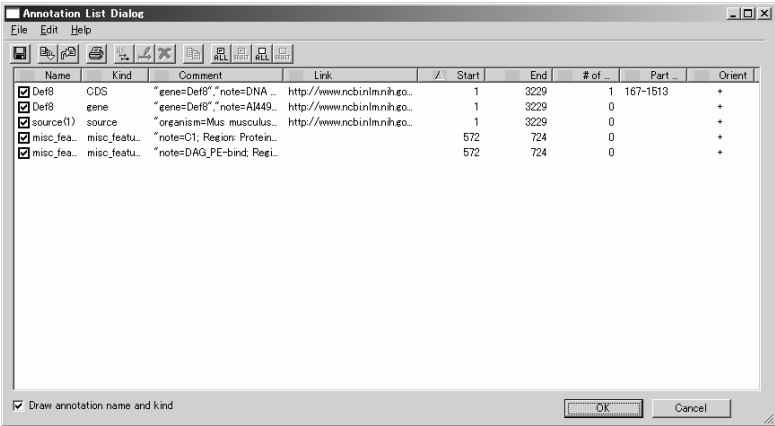
File menu	Description
Save All As...	Stores all the annotations as tab delimited text.
Save Selected As...	Stores the selected annotations as tab delimited text.
Import Annotation	Imports annotations.
Export All Annotation	Exports all the annotations.
Export Selected Annotation	Exports the selected annotations.
Print Setup...	Sets the paper size to use for printing.

Print	Starts printing.

Edit menu	**Description**
New Annotation	Adds new annotations. Opens the Annotation Setting dialog.
Edit Annotation	Edits the selected annotation (single selection only). Opens the Annotation Setting dialog.
Delete Annotation	Deletes the selected annotations.
Copy All	Copies all the annotations to the clipboard as tab-delimited text (with headers).
Copy Selected	Copies the selected annotations to the clipboard as tab-delimited text (with headers).
Check All	Checks all the annotations.
Check Selected	Checks the selected annotations.
Uncheck All	Unchecks all the annotations.
Uncheck Selected	Unchecks the selected annotations.
Select All	Selects all the annotations.

Help menu	**Description**
Help	Displays online help.

Toolbar

Icon	**Description**
	The same as Save As All... from File in the menu.
	The same as Import Annotation from File in the menu.
	The same as the Export Annotation from File in the menu.
	The same as the Print from File in the menu.
	The same as New Annotation from Edit in the menu.
	The same as Edit Annotation from Edit in the menu.
	The same as Delete Annotation from Edit in the menu.
	The same as Copy Selected from Edit in the menu.
	The same as Check All from Edit in the menu.
	The same as Check Selected from Edit in the menu.
	The same as Uncheck All from Edit in the menu.
	The same as Uncheck Selected from Edit in the menu.

Chapter 5 Databases

5.1 List of Databases

The following table lists the databases that DNASIS provides.   These databases are listed under "Database" in the analysis tools menu.   For details about each database, see the page shown in the Page column.

Database name	Page
Sequence database	227
In-house database registration (DNA sequence)	230
In-house database registration (amino acid sequence)	230
Vector database	232
Amino acid motif database	239
Restriction enzyme database	243
Multiple alignment profile	250
Codon table	252
DNA motif database	253
Proteolytic enzyme database	257
BLAST search dedicated database	262

5.2 Sequence Database

The Sequence Database Manager lets you manage sequence databases.   For example, it lets you create, delete, or browse a sequence database.   You can use a sequence database to create a database for BLAST searches or Smith-Waterman searches.
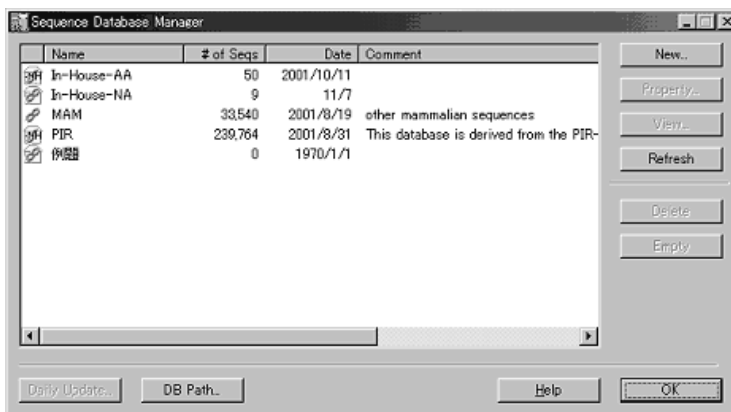
Item	Description
	This icon indicates that the database stores DNA sequence data (GenBank or FASTA files) provided by NCBI.
	This icon indicates that the database stores amino acid sequence data (GenBank or FASTA files) provided by NCBI.
	This icon indicates that the database stores in-house DNA sequence data (such as the experimental data available).
	This icon indicates that the database stores in-house amino acid sequence data (such as the experimental data available).
Name	Displays the name of the database.
# of Seqs	Displays the number of sequence data items stored in the database.
Date	Displays the date on which the database was updated last.
Comment	Displays comments, if any.
New...	Creates a new database. Clicking this button displays the New Database dialog box.
Property...	Displays information about the selected database.
View...	Shows the entries registered in the selected databases.
Refresh	Updates the database list with the latest information.
Delete	Deletes the selected database.
Empty	Deletes all entries from the selected database. You can use this button, for example, if you have inadvertently registered a large number of sequences in an in-house database.
Daily Update...	Not supported in this version.
DB Path...	Allows you to set the path of the directory to store the database. Usually, you do not need to modify the path.
Help	Displays online help.
OK	Exits from the Sequence Database Manager.

Creating a New Database

Click New... in the DNA Sequence Database Manager screen.   The New Database dialog box appears.   A database will be created with the settings specified here.   You can specify the type of the database that the Sequence Database Manager will create by selecting either the Public or In-house tab.   When the dialog box is opened, it displays the Public tab.   Details follow.
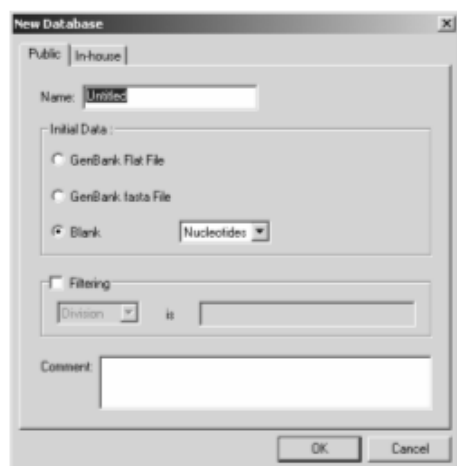
Item	Description
Public	Creates a database to store GenBank or other public data (data to which a unique ID is assigned).
In-house	Creates a database to store in-house data (such as the experimental data available).

Public

Creates a database to store GenBank or other public data.   Data you register in this database must have a unique ID assigned.
Selecting the Public tab in the New Database dialog box displays the following dialog box.
The DNASpace option is required to update a public database.
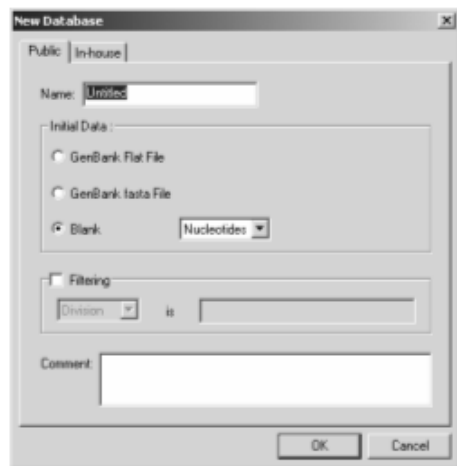
Item	Description
Name	Enter the name of the database.
Initial Data	Select the initial data for the database to be registered.
GenBank Flat File	Read and register data from a GenBank Flat file.
GenBank FASTA File	Read and register data from a GenBank FASTA file.
Blank	Creates an empty database. In the combo box, select either DNA or amino acid sequence database.
Filtering	Optional: Set a filter used to add entries to the database. Filter works only on GenBank format. You can select one of the following four conditions:
Division	Filter by Division field
Definition	Filter by term in Definition field
Keyword	Filter by term in Keyword field
Organism	Filter by Organism field

Comment	Enter a comment. You can leave this field blank.

In-house

Creates a database to store in-house data (such as experimental sequence data available in-house).
Selecting the In-house tab in the New Database dialog box displays the following dialog box.

Item	Description
Name	Enter the name of the database.
Data Source	Set a DNA sequence or amino acid sequence, and initial data.
Blank (Nucleotide)	Create an empty DNA sequence database.
Blank (Amino Acid)	Create an empty amino acid sequence database.
FASTA file (Nucleotides)	Read DNA sequence data in the FASTA format.
FASTA file (Amino Acid)	Read amino acid sequence data in the FASTA format.
Comment	Enter a comment. You can leave this field blank.

5.3 Registering an In-House Database

Register a DNA or amino acid sequence on the sequence editor in an in-house database.

Selecting a Destination Database

1. Click the In-House Database Registration button (either DNA or amino acid) and an Analysis dialog box will appear. Click the Parameter button and a Sequence Database Updater Parameterset Editor will appear.
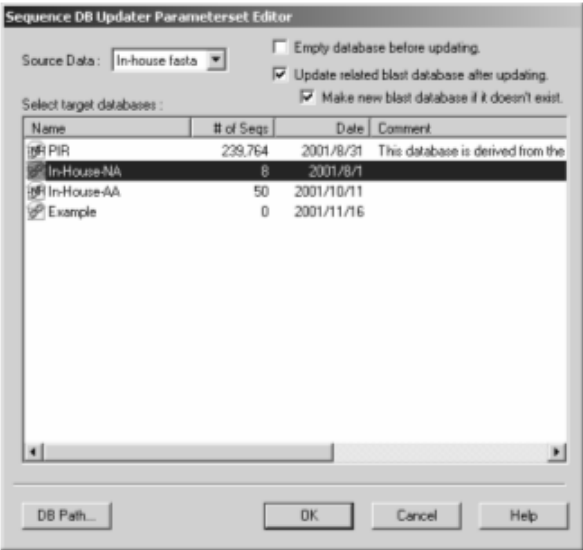
2. The Select target databases field shows a list of databases.   Click to highlight the database to which you want to register a sequence.

3. Select OK.

Registering a Sequence in the Database

1. As explained before, select the database to which you want to register a sequence.

2. Click the sequence on the sequence view.

3. Click the in-house database register button.

Creating an In-house Database

1. In the analysis button view, click the sequence database button.   The DNA Sequence Database Manager appears.

2. Click New... in the DNA Sequence Database Manager screen.   The New Database dialog box appears.

3. Click the in-house tab.

4. Make necessary settings for the database to create, and click the OK button.   The following table describes the items on the screen. .

Name	Enter the name of the database.
Data Source	Set a DNA sequence or amino acid sequence, and initial data.
Blank (Nucleotide)	Create an empty DNA sequence database.
Blank (Amino Acid)	Create an empty amino acid sequence database.
FASTA file (Nucleotides)	Read DNA sequence data in the FASTA format.
FASTA file (Amino Acid)	Read amino acid sequence data in the FASTA format.
Comment	Enter a comment. You can leave this field blank.

Summary of the Parameter Set and Description of Each Parameter
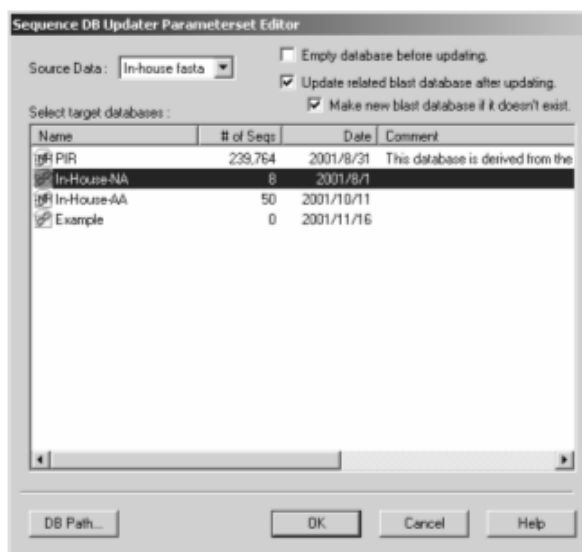
Item (parameter)	Description
Source Data	GenBank FASTA: Not supported by the current version of DNASIS.
	In-house FASTA: Usually, select this item.
Select target databases	Select the database you want to update. You can select one or more databases.
⬦	This icon indicates that the database stores public DNA sequences (mainly GenBank). Each entry in this database has the same unique ID as that used in the original database.
⬦	This icon indicates that the database stores public amino acid sequences.
⬦	This icon indicates that the database stores in-house DNA sequences (such as the experimental data available).
⬦	This icon indicates that the database stores in-house amino acid sequences.
Name	Displays the name of the database.
# of Seqs	Displays the number of sequence data items stored in the database.
Date	Displays the date on which the database was updated last.
Comment	Displays a comment, if any.
DB Path...	Allows you to set the path of the directory to store the database.

5.4 Vector Database

You can list information about the vectors registered in the database.

Window Description
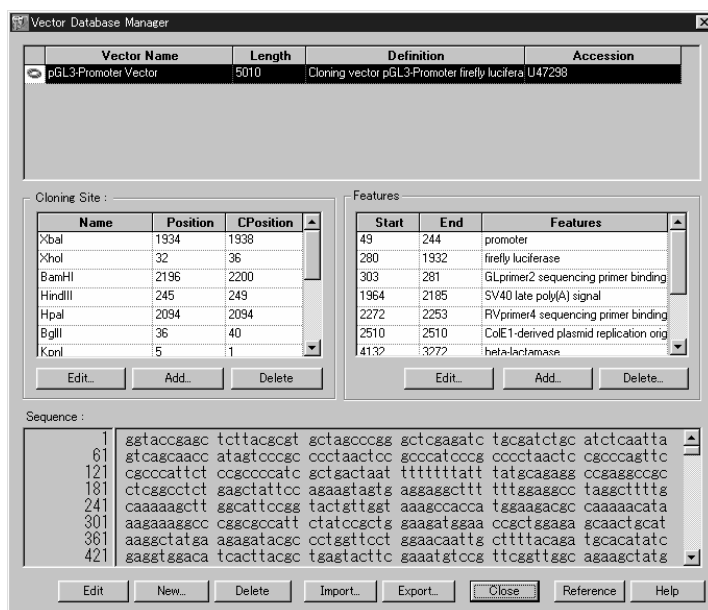
Button	Description
Edit	Modifies the vector.
New	Adds a new vector.
Delete	Deletes the vector selected in the vector list.
Import	Imports a vector from a specified file to the database.
Export	Outputs the vector information selected in the vector list to a file so that DNASIS can import it into another PC.
Close	Closes the Vector Database Manager.
Reference	Displays reference information for the vector selected in the vector list.
Help	Displays online help.

Cloning Site
Specify cloning site settings for a vector.

Button	Description
Edit	Displays the screen used to update the cloning site selected in the cloning site list.
Add	Displays the screen used to add a new cloning site.
Delete	Deletes the cloning site selected in the cloning site list.

Features
Specify feature settings for a vector.

Button	Description
Edit	Displays the screen used to update the features selected in the features list.
Add	Displays the screen used to add a new feature.
Delete	Deletes the feature selected in the features list.

Creating a New Vector

To add a new vector, perform the following steps:

1. In the Vector Database Manager, click the New button.

2. The New Vector dialog box appears.

You can also import a sequence from an external definition file*.

3. Specify the required items in the New Vector dialog box. (See Table 1 for the required items.)

Table 1

Item	Required
Enter New Vector Name	x
Type of Vector	x
Sequence	x
Definition	
Accession	
Features	

*Refer to "Importing a sequence from an External Definition File" in "5.4 Vector Database".

Note: A checkmark (x) indicates a required item.

4. After specifying the required items, click the OK button.

5. The vector list in the Vector Database Manager will display the name of the added vector.

Vector sequence

Table 2 lists the characters you can register as a vector sequence.
You can only register a character that is defined in Table 2.

A	a	M	M
C	c	S	S
G	g	W	W
T	t	B	B
U	u	D	D
R	r	H	H
Y	y	V	V
K	k	N	n

Table 2

Modifying Vector Information

To update information about a vector, perform the following steps:

1. Select the vector you want to update from the vector list in the Vector Database Manager.

2. Modify information as required.

The following describes the procedure for modifying different types of information:
Modifying the definition, accession, vector name, or vector type

1. Select the vector from the vector list and click the Edit button.

2. The Vector dialog box appears.

3. Each item in the Vector dialog box displays the current information.

4. Modify information as required and click the OK button. (If you do not want to modify information, click Cancel.)
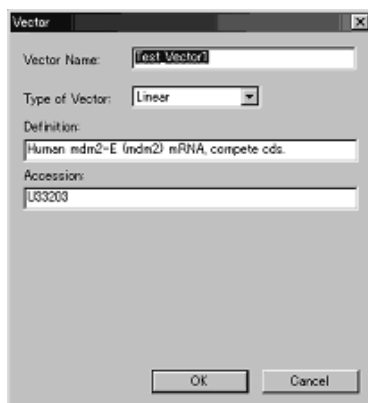
Modifying a Cloning Site

You can add, modify, or delete a cloning site.

Adding a cloning site

To add a new cloning site to the selected vector, perform the following steps:

1. Click the Add button for the Cloning Site.

2. The Cloning Site dialog box appears.
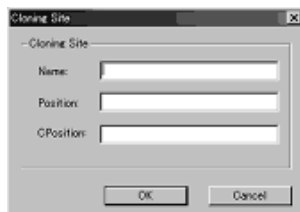
3. In the Cloning Site dialog box, set the Name, Position, and CPosition, and click the OK button.

Note: If you click the OK button without setting the CPosition, the Position setting is automatically copied to the CPosition.

4. Once a cloning site has been added, the cursor moves to the added site.

Modifying a cloning site

To modify a cloning site registered with the selected vector, perform the following steps:
1. Select the cloning site from the list and click the Edit button for the Cloning Site.
2. The Cloning Site dialog box appears.
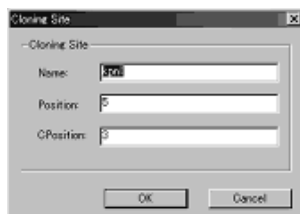
3. Each item in the Cloning Site dialog box displays the current information.
4. In the Cloning Site dialog box, modify the Name, Position, and CPosition, and click the OK button.

Note: If you click the OK button without setting the CPosition, the Position setting is automatically copied to the CPosition.

Deleting a cloning site

To delete a cloning site registered with the selected vector, perform the following steps:
1. Select the cloning site from the list and click the Delete button for the Cloning Site.
2. When a message asking you to confirm deletion appears, select Yes. The selected cloning site is deleted from the database.
3. Once the cloning site has been deleted, the cursor moves to the first site in the list.

Modifying a Feature

You can add, modify, or delete a feature.

Adding a feature

To add a new feature to the selected vector, perform the following steps:

1. Click the Add button for the Features.

2. The Features dialog box appears.

3. In the Features dialog box, set the Name, Start, and End, and click the OK button.

4. Once a feature has been added, the cursor moves to the added feature.

Modifying a feature
To modify a feature registered with the selected vector, perform the following steps:
1. Select the feature from the list and click the Edit button for the Features.
2. The Features dialog box appears.
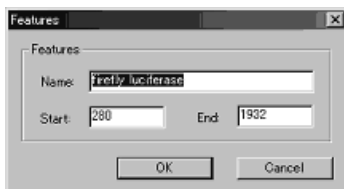
3. Each item in the Features dialog box displays the current information.
4. In the Features dialog box, modify the Name, Start, and End, and click the OK button.

Deleting a feature
To delete a feature registered with the selected vector, perform the following steps:
1. Select the feature from the list and click the Delete button for the Features.
2. When a message asking you to confirm deletion appears, select Yes. The selected feature is deleted from the database.
3. Once the feature has been deleted, the cursor moves to the first feature in the list.

Deleting a Vector

To delete a vector, perform the following steps:

1. Select the vector you want to delete from the vector list in the vector management window.

2. Click the Delete button.

3. When a message asking you to confirm deletion appears, select Yes. The selected vector is deleted from the database.

Displaying References

In the Vector Database Manager screen, you can click the Reference button to view a list of reference information that is set for the vector.
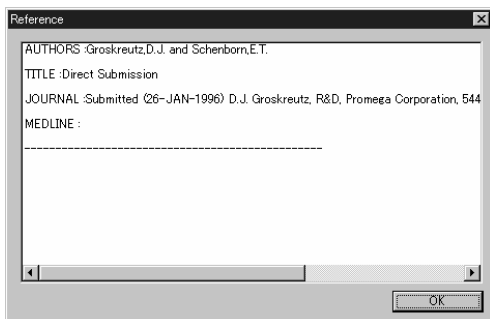
Importing a Sequence from an External Definition File

In the Vector Database Manager, you can create a new vector by importing the contents of an external definition file. To import a sequence, perform the following steps:

1. In the Vector Database Manager screen, click the New button.

2. The New Vector dialog box appears.

3. In the New Vector dialog box, click the Import button.

4. Select the external definition file to import, and click the Open button.

Items imported from an external definition file
You can search the external definition file for several keywords to import the values defined with those keywords.
Table 3 shows relationship between keywords and the items to be imported.
Table 3

Item	Required	Import	Keyword
Enter New Vector Name	x	-	
Type of Vector	x	x	Searches for the external definition file for the keyword LOCUS. Imports the item in the Circular mode if it is defined as Circular; otherwise, imports the item in the Linear mode.
Sequence	x	x	Searches the external definition file for the keyword ORIGIN. Imports the lines up to the // line as the Sequence.
Definition		x	Searches and imports the external definition file for the keyword DEFINITION. You can define more than one DEFINITION.
Accession		x	Searches and imports the external definition file for the keyword ACCESSION.
Features		x	Searches and imports the external definition file for the keyword FEATURES. You can define more than one FEATURES.
Reference		x	Searches the external definition file for the keyword REFERENCE. You can define more than one REFERENCE.

Files you can import

File format
You can import files complying with the GenBank format.

How to import a file
Search a file that complies with the GenBank format for primary search keys.   If the key is found in the file, the value defined with that key will be imported.
You must define primary search keys in the order in which they are shown in the following table:

Primary search key	Secondary search key	Description
LOCUS		Searches for the string "Circular" and, if found, recognizes the item as Circular.
DEFINITION		Imports a string, excluding the string "DEFINITION" itself, as the definition.
ACCESSION		Imports a string, excluding the string "ACCESSION" itself, as the accession.
REFERENCE		Regards the string "REFERENCE" as the start of a reference. Until DNASIS finds a next "REFERENCE" or finds a "FEATURES", it searches for AUTHORS, TITLE, JOURNAL, and MEDLINE as the definitions for a single REFERENCE item.

Primary search key	Secondary search key	Description
	AUTHORS	Imports a string, excluding the string "AUTHORS" itself, as the authors.
	TITLE	Imports a string, excluding the string "TITLE" itself, as the title.
	JOURNAL	Imports a string, excluding the string "JOURNAL" itself, as the journal.
	MEDLINE	Imports a string, excluding the string "MEDLINE" itself, as medline.
FEATURES	See Table 4.	Regards the string "FEATURES" as the start of FEATURES. Searches for FEATURES as a secondary search key. If the Features Key is CDS, imports the Product definition as the Feature name, Start, and End. Otherwise, imports the note definition as the Feature name, Start, and End. If no definition is found, imports the Features Key (see Table 4) as the Feature name and import Start and End as blank*.
	ORIGIN	Regards the string "ORIGIN" as the start of the vector sequence. Import the lines up to the // line as the vector sequence.
Features Key (Table 4)	CDS	
	TATA_signal	
	CAAT_signal	
	promoter	
	enhancer	
	rep_origin	
	polyA_signal	
	primer_bind	
	misc_binding	

*Refer to "Defining Start and End" in " Importing a Sequence from an External Definition File" of "5.4 Vector Database".

Defining Start and End

You cannot import the definition of a join by setting the Start and End positions defined with the Features key.
You can import only the following definitions:

    Features Key   1000 - 1100
    Features Key   complement(1000 - 1100)

The following describes the sections that will be imported, using an example with a GenBank file.
The search keys are shown in boldface type.
The sections to be imported are shown in italics with underlines.

```
LOCUS         HSU33203 309 bp mRNA PRI 20-SEP-1995
DEFINITION    Human mdm2-E (mdm2) mRNA, complete cds.
ACCESSION     U33203
NID           g992684
KEYWORDS      .
SOURCE        human.
ORGANISM      Homo sapiens
              Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
              Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (sites)
AUTHORS       Sigalas,I. and Lunec,J.
TITLE         Multiple alternate spliced mdm2 transcripts with loss of p53
              binding domain sequences: transforming ability and frequent
              detection in human cancer
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 309)
AUTHORS       Lunec,J.
TITLE         Direct Submission
JOURNAL       Submitted (04-AUG-1995) John Lunec, Cancer Research Unit,
              University of Newcastle Upon Tyne - Medical School,
              Framlington Place, Newcastle Upon Tyne, NE2 4HH, U.K
FEATURES      Location/Qualifiers
Source        1..309
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /map="12q"
              /sex="female"
              /tissue_type="primary ovarian tumor"
              /chromosome="12"
gene          1..309
```

/gene="mdm2"

CDS   *1..309*
/gene="mdm2"
/note="mdm2 alternatively spliced form (e)"
/codon_start=1
/evidence=experimental
/product="*mdm2-E*"
/db_xref="PID:g992685"
/translation="MCNTNMSVPTDGAVTTSQIPASEQETLVRPK
PLLLKLLKSVGAQKDTYTMKEVLFYLGQYIMTKRLYD EK
QQHIVNDCANLFPLVDLSIRELYISNYITLGI"

BASE COUNT 100 a 57 c 53 g 99 t
ORIGIN
1   *atgtgcaata ccaacatgtc tgtacctact gatggtgctg taaccacctc acagattcca*
61   *gcttcggaac aagagaccct ggttagacca aagccattgc ttttgaagtt attaaagtct*
121   *gttggtgcac aaaaagacac ttatactatg aaagaggttc ttttttatct tggccagtat*
181   *attatgacta aacgattata tgatgagaag caacaacata ttgtaaatga ttgtgctaac*
241   *ttatttcccc tagttgacct gtctataaga gaattatata tttctaacta tataacccta*
301   *ggaatttag*
//

Importing a Vector

You can add a vector by importing a vector information file exported from DNASIS MAX on another PC.   In the Vector Database Manager, click the Import button and specify a file.
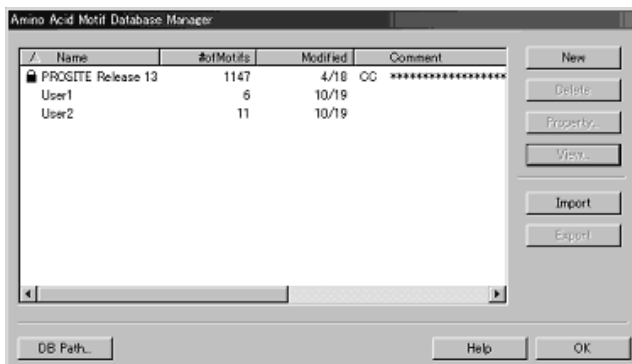
Exporting a Vector

You can output vector information into a file so that you can import a created vector into DNASIS MAX on another PC. In the Vector Database Manager, click the Export button and save vector information to a file.

5.5 Amino Acid Motif Database

The Amino Acid Motif Database Manager lets you browse and manipulate a motif database for amino acid sequences as well as browse and manipulate motif data.
You can create, edit, delete, import, and export a motif database.

Window Description

Item	Description
Motif Database	Displays a list of amino acid motif databases.
Name	Displays the name of the amino acid motif database. Any locked database is shown with a padlock icon to the left of the name. You can click the column header to sort the databases by name.
# of Motifs	Displays the number of motifs registered in the amino acid motif database. You can click the column header to sort the databases by number of motifs.
Modified	Displays the date on which the amino acid motif database was modified last. You can click the column header to sort the databases by date.
Comment	Displays comments for the amino acid motif database, if any. You can click the column header to sort the databases by comments.
New button	Creates a new database. Clicking this button causes a new empty database to be created and added to the list.
Delete button	Deletes the selected motif database. This button is disabled if no database is selected. You cannot delete a locked database.
Property button	Displays the properties of the selected motif database. This button is disabled if no database is selected or if more than one database is selected. The Amino Acid Motif Database Property dialog box appears.
View button	Displays motif data from the selected motif database. This button is disabled if no database is selected or if more than one database is selected. The Amino Acid Motif Database dialog box appears.
Import button	Imports a motif database from an external file. Clicking this button opens a file dialog box that lets you select the motif database to import. DNASIS does not import a motif database if it already contains a database having the same name.
Export button	Exports the selected motif database. This button is disabled if no database is selected or if more than one database is selected. Clicking this button opens a file dialog box that lets you specify where you want to export the motif database to.
DB Path... button	Allows you to specify the location of amino acid motif databases. If the list does not display any registered databases, you can click this button to specify where you want your databases stored. The Amino Acid Motif Database Directory dialog box appears.
Help button	Displays online help.

Editing the Contents of a Motif Database

In the Amino Acid Motif Database Manager, double-click which database you want to display contents for.   The Database Property dialog box appears.
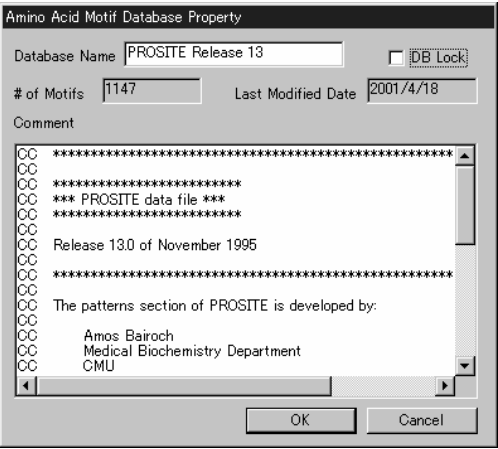
The following describes details about this dialog box:

Item	Description
Database Name	Name of the motif database. The database name must not exceed 64 characters. It cannot contain double-byte characters and these characters that are not supported for file names (/ : , ; * ? " < > |).
DB Lock	Indicates the lock state of the motif database. Place a checkmark in this box to lock the database to prevent it from being edited.
# of Motifs	Displays the number of motifs stored in the database.
Last Modified Date	Displays the date when the database was edited last.
Comment	Displays comments for the database, if any. You can edit the comments if the database is not locked.
OK button	Saves the changes made to the properties and closes the Amino Acid Motif Database Property dialog box. The changes are canceled if a database having the same name is already registered.
Cancel button	Discards the changes made to the properties and closes the Amino Acid Motif Database Property dialog box.

Displaying a List of Registered Amino Acid Motifs

You can display a list of all motifs registered in the amino acid motif database.   In the Amino Acid Motif Database Manager, select the motif database for which you want to display a list and click the View... button.
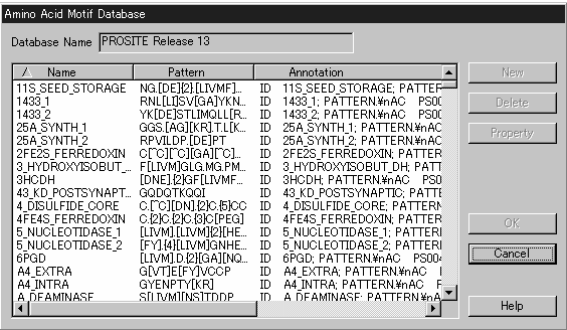

You can create, edit, or delete amino acid motif data.

Item	Description
Database Name	Displays the name of the database.
Motif data list	A list of all motifs in the database. The list shows the name, pattern, and annotation of each motif.
Name column	Displays the name of the motif. Click the column header to sort the motifs by name.
Pattern column	Displays the pattern of the motif. Click the column header to sort the motifs by pattern.
Annotation column	Displays the annotation of the motif. Click the column header to sort the motifs by annotation.
New button	Creates a new motif. This button is disabled if the database is locked. Clicking this button makes the Amino Acid Motif Property dialog box appear.
Delete button	Deletes the selected motif. This button is disabled if the database is locked or no motif is

Item	Description
	selected. You can also delete more than one motif at one time.
Property button	Allows you to edit the selected motif data. This button is disabled if no motif is selected or if more than one motif is selected. Clicking this button makes the Amino Acid Motif Property dialog box appear. If the database is locked, you can browse the motif data but cannot edit it.
OK button	Saves the changes made to the motif data and closes the Amino Acid Motif Database dialog box. This button is disabled if the database is locked.
Cancel button	Discards the changes made to the motif data and closes the Amino Acid Motif Database dialog box.
Help button	Displays online help.

Displaying Motif Properties

From the motif list in the Amino Acid Motif Database dialog box (see the previous section), select the motif for which you want to display properties and click the Property button.

Details follow:

Item	Description
Motif Name	Name of the motif. You cannot edit this item if the database is locked.
	The motif name must not exceed 255 characters.
Motif Pattern	Pattern of the motif. You cannot edit this item if the database is locked.
Pattern Assistant	A drop-down list which helps you specify a motif pattern. You cannot use this list if the database is locked. The available items include:
Beginning of the sequence	Enters a caret character (^) at the beginning of a sequence.
Any character	Enters a period (.), which matches any character.
End of sequence	Enters a dollar sign ($) at the end of a sequence.
Or	Enters a vertical bar (|), which means "or".
Grouping	Enters parentheses ( ) for grouping.
Character Class	Enters brackets [ ], which means a range of characters.
Character not in the list	Enters a caret and a space within brackets [^], which means characters other than those in the specified range.
Match 0 or more times	Enters an asterisk (*), which indicates zero or more repetitions.
Match 1 or more times	Enters a plus sign (+), which indicates one or more repetitions.
Match 0 or 1 times	Enters a question mark (?), which indicates zero or one repetition.
Match exactly n times	Enters braces { }, which means n repetitions.
Match at least n times	Enters a comma {,}, which means n or more repetitions.
Motif Pattern Test Sequence	Enters a sequence used to test the pattern. Clicking the Test button causes any section that matches the pattern to be highlighted.
Test button	This button is used with the Motif Pattern and Motif Pattern Test Sequence fields to test the

Item	Description
	pattern. Clicking the Test button causes any section that matches the Test Sequence pattern to be highlighted. If more than one section matches, only the first match is highlighted. This button is disabled if the Motif Pattern or Motif Pattern Test Sequence is not specified. It is also disabled if the Motif Pattern Test Sequence contains anything other than alphabetic characters.
< button	If more than one section matches as a result of a pattern test, clicking this button highlights the match previous to the one currently highlighted. This button is disabled if the first match is currently highlighted.
> button	If more than one section matches as a result of a pattern test, clicking this button highlights the match following the one currently highlighted. This button is disabled if the last match is currently highlighted.
Motif Annotation	Annotation of the motif. You cannot edit this item if the database is locked.
Help button	Displays online help.
OK button	If you have opened the dialog box from the Property button, the OK button saves the changes made to the motif data and closes the dialog box. If you have opened the dialog box from the New button, the OK button adds the motif data and closes the dialog box. You cannot register a motif having the same name as that of an existing motif. You cannot register a motif if its motif pattern is invalid. This button is disabled if the database is locked.
Cancel button	Discards all changes and closes the dialog box.

Adding a Motif Database

1. In the analysis button view, click an amino acid motif database.

2. Click the New button in the Amino Acid Motif Database Manager, as shown in the figure.   DNASIS creates a database "Untitled" in the window.
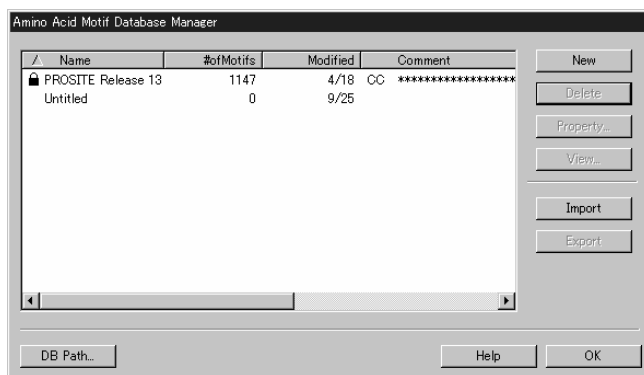
3. Click the "Untitled" database to highlight it.

4. Click the Property... button. The Amino Acid Motif Database Property dialog box appears.   Make necessary settings.

5. Click the OK button.

5.6 Restriction Enzyme Database

You can display a list of restriction enzyme databases.   DNASIS MAX supports the functions for creating, editing, deleting, importing, and exporting restriction enzyme data.

Window Description

The Restriction Enzyme Database Manager consists of the main Restriction Enzyme Database Manager window, the New Enzyme window for creating new restriction enzyme data, and the Enzyme Property window for editing restriction enzyme data.

Restriction Enzyme Database Manager window

This is the main window of the Restriction Enzyme Database Manager.

Parameter Description

Item (parameter)	Description
Name	Name of the restriction enzyme.
Recognition Sequence	Indicates the sequence that the restriction enzyme recognizes in the direction from 5' to 3'. An exclamation mark (!) indicates the position to cut. If the recognition sequence does not have a palindrome structure, the sequence recognized by a Normal sequence and that recognized by a Complementary sequence are separated with a slash (/).
Bases	Indicates the number of bases in the recognition sequence.
Kind Of Cut	Kind Of Cut Indicates the shape of the cut performed by the restriction enzyme.
5'-extended	Cuts the sequence so that the 5' end is longer than the 3' end. 5' - G A A T T C - 3' 3' - C T T A A G - 5'
3'-extended	Cuts the sequence so that the 3' end is longer than the 5' end. 5' - T G C G C A - 3' 3' - A C G C G T – 5'
blunt-cut	Cuts the sequence so that the 3' and 5' ends have the same length. 5' - C C C G G G - 3' 3' - G G G C C C - 5'
not identified	Indicates that you cannot identify the position to cut for this restriction enzyme. If you check this enzyme, it will not be registered as a parameter.

Button	Description
New... button	Creates a new restriction enzyme. The New Enzyme window appears.
Delete button	Deletes the selected restriction enzyme from the database. To delete more than one enzyme, select the enzymes you want to delete and then click this button.
Property... button	Allows you to edit the restriction enzyme. The Enzyme Property window appears.
Import... button	Imports restriction enzyme data from an external file.
Export... button	Exports restriction enzyme data to an external file.
Help button	Displays online help.
OK button	OK button Exits from the Restriction Enzyme Database Manager.

New Enzyme window

You can use this window to create a new restriction enzyme.

Details follow:

| Item (parameter)	Description
Enzyme Name text area	A text area used to enter the name of the restriction enzyme you want to create. The OK button is disabled if this area is blank.
Normal text area	A text area used to enter the recognition sequence for the restriction enzyme you want to create. Enter the recognition sequence using a complex code containing characters from the string ACGTURYWSKMBDHVN (not case sensitive). Enter an exclamation mark (!) at the position to cut. Specify the characters in the direction from 5' to 3'.
Complementary text area	A text area used to enter a complement recognition sequence if the restriction enzyme you create does not have a palindrome-structured recognition sequence. Enter the recognition sequence using a complex code containing characters from the string ACGTURYWSKMBDHVN (not case sensitive). Enter an exclamation mark (!) at the position to cut. Specify the characters in the direction from 5' to 3'.
Text area for the number of bases	Automatically filled with the same length value of the Normal recognition sequence, excluding an exclamation mark (!). If you enter a recognition sequence in the Complementary area, it must have the same length, excluding the exclamation mark, as that specified here.
Combo box for the Kind of Cut	Automatically selects the cut shape for the restriction enzyme.
OK button	You can click the OK button to register the new restriction enzyme created. This button is disabled if DNASIS detects any of the following errors in the data you have entered: 1. The Enzyme Name text area does not contain a restriction enzyme name. 2. The Normal text area does not contain a recognition sequence. 3. The Normal text area contains a character other than ACGTURYWSKMBDHVN and !. 4. The Normal text area does not contain an exclamation mark (!) or it contains more than one exclamation mark. 5. The Complementary text area contains a recognition sequence including a character other than ACGTURYWSKMBDHVN and !. 6. The Complementary text area contains a recognition sequence without an exclamation mark (!) or more than one exclamation mark. 7. The Complementary text area contains a recognition sequence having a length different from

Item (parameter)	Description
	that of the sequence in the Normal text area.
	When you click the OK button, DNASIS checks for a duplicate enzyme name. If the database already contains a restriction enzyme having the same name, DNASIS shows a dialog box with a message stating that the restriction enzyme name is a duplicate and you cannot register the restriction enzyme.
Cancel button	Cancels the creation of a new restriction enzyme and returns to the Restriction Enzyme Database Manager window.

Example of Registering a Restriction Enzyme

[For EcoR I]

The recognition sequence has a palindrome structure.   Enter G!AATTC in the Normal text area.

```
5'- G|A A T T C - 3'
3'- C T T A A|G - 5'
```

[For Mbo II]

```
5'- G A A G A N N N N N N N N|- 3'
3'- C T T C T N N N N N N N|N - 5'
```

The recognition sequence does not have a palindrome structure.   Enter GAAGANNNNNNNN! in the Normal text area and N!NNNNNNNTCTTC in the Complementary text area.

Enzyme Property Window

You can use this window to edit a restriction enzyme.

Item (parameter)	Description
Enzyme Name text area	A text area used to enter the name of the restriction enzyme you want to create. The OK button is disabled if this area is blank.
Normal text area	A text area used to enter the recognition sequence for the restriction enzyme you want to create. Enter the recognition sequence using a complex code containing characters from the string ACGTURYWSKMBDHVN (not case sensitive). Enter an exclamation mark (!) at the position to cut. Specify the characters in the direction from 5' to 3'.
Complementary text area	A text area used to enter a complement recognition sequence if the restriction enzyme you create does not have a palindrome-structured recognition sequence. Enter the recognition sequence using a complex code containing characters from the string ACGTURYWSKMBDHVN (not case sensitive). Enter an exclamation mark (!) at the position to cut. Specify the characters in the direction from 5' to 3'.
Text area for the number of bases	Automatically filled with the same length value of the Normal recognition sequence, excluding an exclamation mark (!). If you enter a recognition sequence in the Complementary area, it must have the same length, excluding the exclamation mark, as that specified here.
Combo box for the Kind of Cut	Automatically selects the cut shape for the restriction enzyme. For details about the cut shape, refer to the description of the Restriction Enzyme Database Manager window.
OK button	You can click the OK button to register the new restriction enzyme created. This button is disabled if DNASIS detects any of the following errors in the data you have entered:
	1. The Enzyme Name text area does not contain a restriction enzyme name.
	2. The Normal text area does not contain a recognition sequence.
	3. The Normal text area contains a character other than ACGTURYWSKMBDHVN and !.
	4. The Normal text area does not contain an exclamation mark (!) or it contains more than one exclamation mark.

| Item (parameter)	Description
	5. The Complementary text area contains a recognition sequence including a character other than ACGTURYWSKMBDHVN and !.
	6. The Complementary text area contains a recognition sequence without an exclamation mark (!) or more than one exclamation mark.
	7. The Complementary text area contains a recognition sequence having a length different from that of the sequence in the Normal text area.
	When you click the OK button, DNASIS checks for a duplicate enzyme name. If the database already contains a restriction enzyme having the same name, DNASIS shows a dialog box with a message stating that the restriction enzyme name is a duplicate and you cannot register the restriction enzyme.
Cancel button	Cancels the creation of a new restriction enzyme and returns to the Restriction Enzyme Database Manager window.

Importing Restriction Enzyme Data

Import file dialog box

In the Restriction Enzyme Database Manager window, clicking the Import button causes the following dialog box to appear:
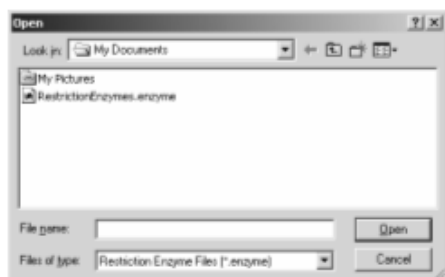
Select a file storing restriction enzyme data and click the Open button.
DNASIS imports restriction enzyme data while performing format checks.

Format Error dialog box

If the restriction enzyme being imported contains a format error, the Format Error dialog box appears.   DNASIS cannot import a restriction enzyme which contains a format error.
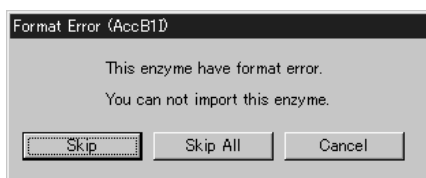
The following describes the conditions which cause a format error.   Check the format of the data you are importing.

Conditions causing a format error

1. The Normal text area does not contain a recognition sequence.
2. The Normal text area contains a character other than ACGTURYWSKMBDHVN and !.
3. The Normal text area does not contain an exclamation mark (!) or it contains more than one exclamation mark.
4. The Complementary text area contains a recognition sequence including a character other than ACGTURYWSKMBDHVN and !.
5. The Complementary text area contains a recognition sequence without an exclamation mark (!) or more than one exclamation mark.
6. The Complementary text area contains a recognition sequence having a length different from that of the sequence in the Normal text area.

Button description

Button	Description
Skip button	Skips the restriction enzyme which contains a format error and continues importing subsequent restriction enzyme data.
Skip All button	Continues importing restriction enzyme data, skipping all subsequent restriction enzymes which contain a format error.
Cancel button	Cancels the importing of restriction enzyme data. Clicking this button preserves the original restriction

enzyme data without importing any data.

Name Collision dialog box

If the database already contains a restriction enzyme having the same name as that of the restriction enzyme being imported, the Name Collision dialog box appears, as shown below.
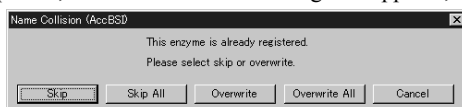
Button description

Button	Description
Skip button	Skips the restriction enzyme having a duplicate name and continues importing subsequent restriction enzyme data.
Skip All button	Continues importing restriction enzyme data, skipping all subsequent restriction enzymes having a duplicate name.
Overwrite button	Overwrites the existing restriction enzyme with the imported one and continues importing subsequent restriction enzyme data.
Overwrite All button	Continues importing restriction enzyme data, overwriting all existing restriction enzymes with the imported ones.
Cancel button	Cancels the importing of restriction enzyme data. Clicking this button preserves the original restriction enzyme data without importing any data.

Non-Enzyme datum window

This dialog box appears if the data being imported contains any data other than restriction enzyme data.
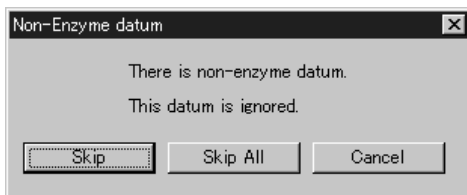
DNASIS cannot import a restriction enzyme which contains this error.   Check the format of the data you are importing.

Button	Description
Skip button	Skips the non-restriction enzyme data and continues importing subsequent restriction enzyme data.
Skip All button	Continues importing restriction enzyme data, skipping all subsequent non-restriction enzyme data.
Cancel button	Cancels the importing of restriction enzyme data. Clicking this button preserves the original restriction enzyme data without importing any data.

Registering a New Restriction Enzyme

1. Click the restriction enzyme database button to open the Restriction Enzyme Database Manager.
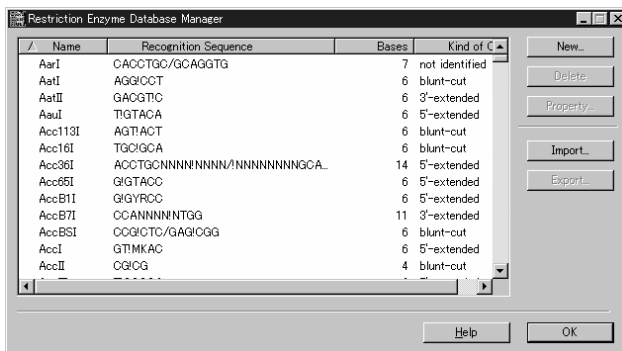
2. Click the New... button.   The New Enzyme dialog box appears, as shown below.   Enter necessary information and click OK.
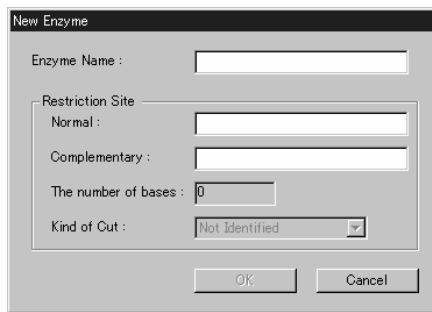
Exporting a Restriction Enzyme

You can export a selected restriction enzyme using the Restriction Enzyme Database Manager window.

Export file dialog box

In the Restriction Enzyme Database Manager window, clicking the Export button causes the following window to appear:
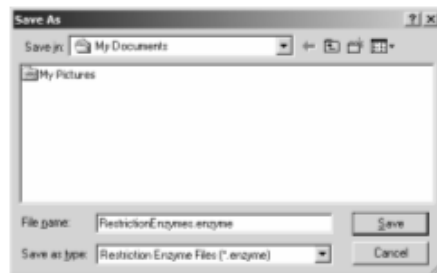
*For details about the format of the output restriction enzyme data, refer to "Restriction Enzyme Data Format" in "5.6 Restriction Enzyme Database".

Specify the destination folder and file name, and click the Save button.   DNASIS outputs the data of the restriction enzyme that was selected when you clicked the Export button.*

Export errors

The following two errors may occur during export:

1. [Enzyme Name] has too long name. You can't export this enzyme.

The name of the restriction enzyme [Enzyme Name] is too long.   DNASIS cannot export a restriction enzyme which contains this error.   Correct the name of the restriction enzyme so that it does not exceed 255 characters and reexport it.

2. You can't export no name enzyme.

The restriction enzyme being exported does not have a name.   DNASIS cannot export a restriction enzyme with this error.   Name the restriction enzyme and reexport it.

Complex Code

Code list
(Not case-sensitive)

Complex code	ACGTT
A	A
C	C
G	G
T, U	T or U
R	A or G
Y	C, T, or U
W	A, T, or U
S	C or G
K	G, T, or U
M	A or C
B	C, G, T, or U
D	A, G, T, or U
H	A, C, T, or U

Complex code	ACGTT
V	A, C, or G
N	A, C, G, T, or U

Restriction Enzyme Data Format

The restriction enzyme data you import or export must be plain text written in the following format.

To describe more than one restriction enzyme, enter a carriage return before describing a next restriction enzyme.

[HSK_REnzymeDB XXXX]	XXXX is the restriction enzyme name (space not allowed; within 255 characters).
NAME=XXXX	Restriction enzyme name (space allowed).
SITE_N=NN!NNN	Normal recognition sequence (cut at !)
SITE_C=	Complementary recognition sequence (described for non-palindrome structure)
	Carriage return
[HSK_REnzymeDB XXXX]	XXXX is the restriction enzyme name (space not allowed; within 255 characters).
NAME=XXXX	Restriction enzyme name (space allowed).
SITE_N=NN!NNN	Normal recognition sequence (cut at !)
SITE_C=	Complementary recognition sequence (described for non-palindrome structure)

5.7 Multiple Alignment Profile

DNASIS MAX supports managing the profile of multiple alignments.   You can create an empty profile, delete a profile, modify the attributes of a profile, import and export a profile.

Multiple alignment profile

What is a profile?
A multiple alignment profile is pre-calculated data for the alignments between multiple input sequences that is saved for later use.

Why do I want to use a profile?
Calculating multiple alignments requires a long time.   DNASIS requires only ten minutes to calculate multiple alignments for 40 data items, but it may require two days for 200 data items.   This applies when the average BP length for the input sequences is about 1.5Kbp.   Longer sequences, such as a gene or a complete genome, require a longer time.

If you have many known sequences and want to calculate alignment between an unknown sequence and the known ones, you can save the time required to calculate alignment with the unknown sequence by creating a profile first. Calculating a profile requires the same time as an ordinary calculation.   However, once a profile is created, DNASIS can calculate alignment with the unknown sequence much faster (in about 10 seconds for the above example).

Disadvantages of using a profile
Using a profile provides fast calculation.   However, it results in degraded alignment precision.   The same data may produce different results when you use a profile and do not use a profile.   You should consider those characteristics when using a profile.

Window Description
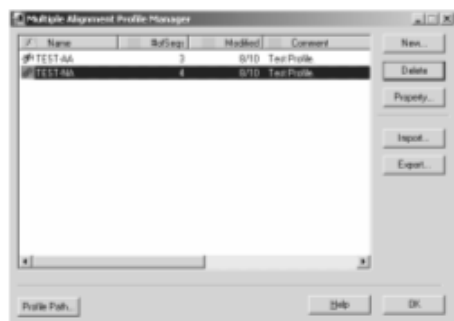
Profile list
Displays a list of multiple alignment profiles.   The following describes the meaning of each column.   You can click the header of the column to sort the list in ascending order using that column as the key.   To sort the list in descending order, click the column header again.

Click a profile to select it.   You can also select a range of profiles by Shift-clicking them and select multiple profiles by Ctrl-clicking them.   You can press the F2 key to edit the profile that currently has a focus (enclosed by dotted lines).

Column name	Description
Name	Displays the name of the profile with an icon indicating the profile type. The following icons are used:
	🧬 : DNA sequence
	-NH: Amino acid sequence
	🔒 : DNA sequence (read-only)
	-NH: Amino acid sequence (read-only)
	Note: Read-only profiles cannot be overwritten during analysis with the multiple alignment profile create button.
# of Seqs	Displays the number of sequences in the profile.
Modified	Displays the date on which the profile was created.
Comment	Displays comments for the profile, if any.

Button name	Description
New...	Creates a new profile.
Delete	Deletes the profile.

Button name	Description
Property...	Displays the attributes of the profile in a dialog box. You can edit some of the attributes.
Import...	Imports a profile from a text file created with the export feature.
Export...	Exports the selected profile or profiles to a text file. Using a text file you created with this feature, you can import the profile to DNASIS running on another machine.
Profile Path...	Allows you to view or modify the path of the directory to store the profile.
Help	Displays online help.
OK	Saves the changes and exits from the Multiple Alignment Profile Manager.

Property Window

In the Multiple Alignment Profile Manager, clicking the Property... button opens this dialog box.
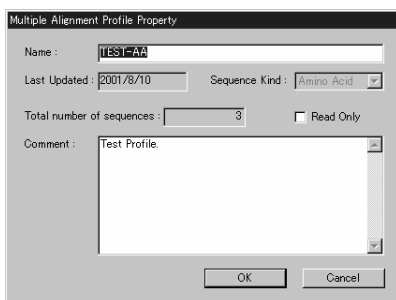
Item	Description
Name	Displays the name of the profile. You can edit the name. You can use up to 64 characters excluding any of the invalid characters (: < > \ / * ? |). You cannot specify the same name as that of an existing profile.
Last Updated	Displays the date on which the profile was updated last.
Sequence Kind	Displays the sequence type (DNA or Protein) of the profile. You can modify this item only when the Total number of sequences is 0.
Total number of sequences	Displays the number of sequences contained in the profile.
Read Only	Check this item if you want to prevent this profile from being overwritten.
Comment	Displays comments for the profile, if any. The comments cannot exceed 32767 characters. You can use any single-byte characters and carriage returns.

5.8 Codon Table

Displays a codon table.   You can edit the contents of a codon table.

Choose a codon table from the Codon Table: list box to display its contents.

Editing a Codon Table

1. Select the codon table you want to edit.

2. Click the ▼ button for the codon to edit.

3. Select the corresponding amino acid.   Select *** for a stop codon.

4. Click the OK button.

Note: User1 to User4 in the Codon Table list are provided for creating a new codon table.

5.9 DNA Motif Database

You can display a list of DNA motif databases.
DNASIS MAX supports the functions for creating, editing, deleting, importing, and exporting DNA motif databases.

Window Description
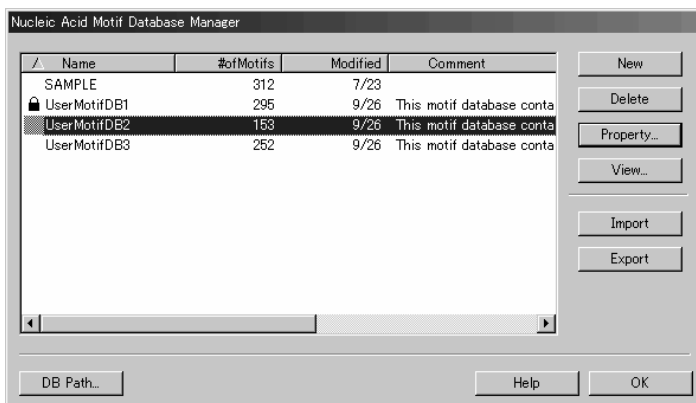
Item	Description
Motif Database list	Displays a list of DNA motif databases.
Name column	Displays the name of the DNA motif database. Any locked database shows a padlock icon to the left of its name.
	You can click the column header to sort the databases by name.
# of Motifs column	Displays the number of motifs registered in the DNA motif database. You can click the column header to sort the databases by number of motifs.
Modified column	Displays the date when the DNA motif database was modified last.
	You can click the column header to sort the databases by date.
Comment column	Displays comments for the DNA motif database, if any.
	You can click the column header to sort the databases by comments.
New button	Creates a new database. Clicking this button causes a new empty database to be created and added to the list.
Delete button	Deletes the selected motif database. This button is disabled if no database is selected. You cannot delete a locked database.
Property button	Displays the properties of the selected motif database. This button is disabled if no database is selected or if more than one database is selected. The Nucleic Acid Motif Database Property dialog box appears.
View button	Displays motif data from the selected motif database. This button is disabled if no database is selected or if more than one database is selected. The Nucleic Acid Motif Database dialog box appears.
Import button	Imports a motif database from an external file.
	Clicking this button opens a file dialog box that lets you select the motif database to import. DNASIS MAX does not import a motif database if it already contains a database having the same name.
Export button	Exports the selected motif database. This button is disabled if no database is selected or if more than one database is selected.
	Clicking this button opens a file dialog box that lets you specify where you want to export the motif database to.
DB Path... button	Allows you to specify the location of nucleic acid motif databases.
	If the list does not display any registered databases, you can click this button to specify where you want your databases stored.
	The Nucleic Acid Motif Database Directory dialog box appears.
Help button	Displays online help.

Editing the Properties of a Motif Database

1. In the analysis button view, click the DNA motif database.

2. The Nucleic Acid Motif Database Property dialog box appears, as shown in the figure.   You can use this dialog box to view and edit the properties of a DNA motif database.   The following describes details about the dialog box:

Item	Description
Database Name	Name of the motif database.
	The database name must not exceed 64 characters. It cannot contain double-byte characters
	and these characters that are not supported for file names (\ / : , ; * ? " < > |).
DB Lock	Indicates the lock state of the motif database. Place a checkmark in this box to lock the
	database to prevent it from being edited.
# of Motifs	Displays the number of motifs stored in the database.
Last Modified Date	Displays the date when the database was edited last.
Comment	Displays comments for the database, if any. You can edit the comments if the database is not
	locked.
OK button	Saves the changes made to the properties and closes the Nucleic Acid Motif Database
	Property dialog box. The changes are canceled if a database having the same name is
	already registered.
Cancel button	Discards the changes made to the properties and closes the Nucleic Acid Motif Database
	Property dialog box.

Displaying a List of Registered DNA Motifs

You can display a list of all motifs registered in the DNA motif database.   In the Nucleic Acid Database Manager, select the database for which you want to list the contents, and click the View... button.
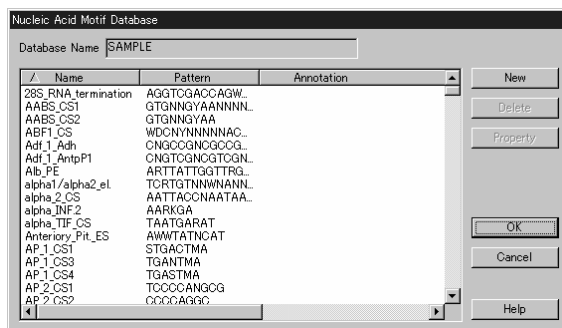
Item	Description
Database Name	Displays the name of the database and a list below of all motifs in the database. The list shows
	the name, pattern, and annotation of each motif.
Name column	Displays the name of the motif. Click the column header to sort the motifs by name.
Pattern column	Displays the pattern of the motif. Click the column header to sort the motifs by pattern.
Annotation column	Displays the annotation of the motif. Click the column header to sort the motifs by annotation.
New button	Creates a new motif. This button is disabled if the database is locked. Clicking this button
	makes the Nucleic Acid Motif Property dialog box appear.
Delete button	Deletes the selected motif. This button is disabled if the database is locked or if no motif is
	selected. You can also delete more than one motif at one time.
Property button	Allows you to edit the selected motif data.

	This button is disabled if no motif is selected or if more than one motif is selected. Clicking this button makes the Nucleic Acid Motif Property dialog box appear. If the database is locked, you can browse motif data but cannot edit it.
OK button	Saves the changes made to the motif data and closes the Nucleic Acid Motif Database dialog box. This button is disabled if the database is locked.
Cancel button	Discards the changes made to the motif data and closes the Nucleic Acid Motif Database dialog box.
Help button	Displays online help.

Editing the Properties of a Motif

You can view and edit the properties of a DNA motif registered in the database.

1. In the analysis button view, click the DNA motif database.   The Nucleic Acid Motif Database Manager appears.

2. Select the database containing the motif you want to display, and click the View... button.   The Nucleic Acid Motif Database dialog box appears.

3. Select the motif you want to edit from the list, and click the Property button. The Nucleic Acid Motif Property dialog box appears, as shown in the figure.   Make the necessary settings.   After completing editing the properties, click the OK button.   The following describes details about the dialog box:

Item	Description
Motif Name	Name of the motif. You cannot edit this item if the database is locked. The motif name must not exceed 255 characters.
Motif Pattern	Pattern of the motif. You cannot edit this item if the database is locked.
Pattern Assistant	A drop-down list which helps you specify a motif pattern. You cannot use this list if the database is locked. The available items include:
Beginning of the sequence	Enters a caret character (^) at the beginning of a sequence.
Any character	Enters a period (.), which matches any character.
End of sequence	Enters a dollar sign ($) at the end of a sequence.
Or	Enters a vertical bar (|), which means "or".
Grouping	Enters parentheses ( ) for grouping.
Character Class	Enters brackets [ ], which means a range of characters.
Character not in the list	Enters a caret and a space within brackets [^], which means characters other than those in the specified range.
Match 0 or more times	Enters an asterisk (*), which indicates zero or more repetitions.
Match 1 or more times	Enters a plus sign (+), which indicates one or more repetitions.
Match 0 or 1 times	Enters a question mark (?), which indicates zero or one repetition.
Match exactly n times	Enters braces { }, which means n repetitions.

Item	Description
Match at least n times	Enters a comma {,}, which means n or more repetitions.
Motif Pattern Test Sequence	Enters a sequence used to test the pattern. Clicking the Test button causes any section that matches the pattern to be highlighted.
Test button	This button is used with the Motif Pattern and Motif Pattern Test Sequence fields to test the pattern. Clicking the Test button causes any section that matches the Test Sequence pattern to be highlighted. If more than one section matches, only the first match is highlighted. This button is disabled if the Motif Pattern or Motif Pattern Test Sequence is not specified. It is also disabled if the Motif Pattern Test Sequence contains anything other than alphabetic characters.
< button	If more than one section matches as a result of a pattern test, clicking this button highlights the match previous to the one currently highlighted. This button is disabled if the first match is currently highlighted.
> button	If more than one section matches as a result of a pattern test, clicking this button highlights the match following the one currently highlighted. This button is disabled if the last match is currently highlighted.
Motif Annotation	Annotation of the motif. You cannot edit this item if the database is locked.
Help button	Displays online help.
OK button	If you have opened the dialog box from the Property button, the OK button saves the changes made to the motif data and closes the dialog box. If you have opened the dialog box from the New button, the OK button adds the motif data and closes the dialog box. You cannot register a motif having the same name as that of an existing motif. You cannot register a motif if its motif pattern is invalid. This button is disabled if the database is locked.
Cancel button	Discards all changes and closes the dialog box.

5.10 Proteolytic Enzyme Database

DNASIS MAX supports the functions for creating, editing, deleting, importing, and exporting proteolytic enzyme data.

Window Description

The window displays a list of proteolytic enzyme data registered in the database.
You can select one or more proteolytic enzymes and manipulate the data.
You can click the column header to sort the data in ascending (□ button) or descending (□ button) order for that column.   Initially, data is displayed in ascending order of the enzyme name.
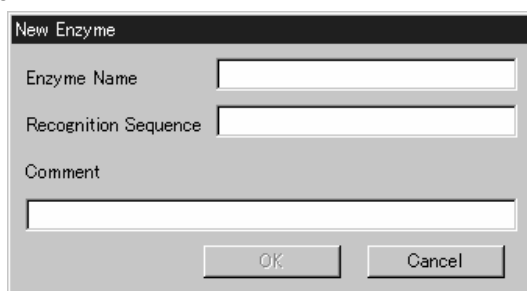
Item (parameter)	Description
Name(NAME)	The name of the registered proteolytic enzyme.
Recognition Sequence(SITE)	Sequence recognized by the proteolytic enzyme.
	An amino acid sequence is represented in the single-character format with an exclamation mark
	(!) indicating a cut position.
	If there is more than one recognition sequence, a slash (/) is used as a delimiter.
	If there is more than one recognition amino acid (complex code), each is enclosed by brackets [ ].
	X indicates any amino acid.
	Example: [KR] ! X / AR ! X
	Identify KX, RX, and ARX and cut between K and X, R and X, and AR and X.
Comment	Displays comments for the proteolytic enzyme, if any.
New... button	Creates new proteolytic enzyme data. The New Enzyme dialog box appears.
Delete button	Deletes all selected proteolytic enzymes.
Property... button	Lets you edit data for the selected proteolytic enzyme.
	The Enzyme Property dialog box appears. This button is disabled if no enzyme is selected or more than one enzyme is selected.
Import... button	Imports exported data for a proteolytic enzyme.
Export... button	Exports data for the selected proteolytic enzyme.
	This button is disabled if no data is selected.
Help button	Displays online help.
OK button	Exits from the Proteolytic Enzyme Database Manager.

Creating New Proteolytic Enzyme Data

In the Proteolytic Enzyme Database Manager, you can click the New... button to create new proteolytic enzyme data. Clicking the New... button causes the New Enzyme dialog box to appear.   Enter data in this dialog box to create enzyme data.   You cannot register an enzyme having the same name as that of any existing enzyme registered in the database.

New Enzyme dialog box

Item (parameter)	Description
Enzyme Name(NAME)	Enter the name of the proteolytic enzyme you want to register. The OK button is disabled if this field is blank.
Recognition Sequence(SITE)	Sequence recognized by the proteolytic enzyme.
	An amino acid sequence is represented in the single-character format with an exclamation mark (!) indicating a cut position.
	If there is more than one recognition sequence, a slash (/) is used as a delimiter.
	If there is more than one recognition amino acid (complex code), acids are enclosed by brackets [ ].
	X indicates any amino acid.
	Example: [KR] ! X / AR ! X
	Identify KX, RX, and ARX and cut between K and X, R and X, and AR and X.
	The OK button is disabled in the following cases:
	For each sequence separated by a slash (/):
	1. There are more than one cut position (!).
	2. The data does not contain any amino acid characters.
	3. Any character other than A to Z, !, and [ ] is used.
	4. Brackets [ ] are nested in other brackets [ ].
	5. Brackets [ ] are not paired.
	6. Brackets [ ] contain X.
	7. Brackets [ ] contain no characters.
Comment	Displays comments for the proteolytic enzyme, if any.
OK button	Creates a new proteolytic enzyme from the entered data. DNASIS cannot register an enzyme if its name is already used for an existing enzyme. In such a case, you must change the name to register it.
Cancel button	Cancels the creation of new proteolytic enzyme data.

Errors that may occur when creating new data

Duplicate enzyme name

If you specify an already registered name for a new enzyme, the dialog box appears.
Click the OK button to return to the New Enzyme dialog box.   Change the name of the enzyme and retry.
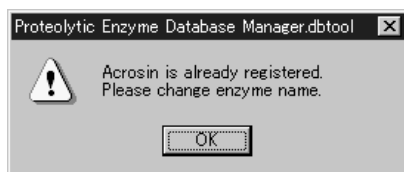
Editing Proteolytic Enzyme Data

In the main window, you can click the Property... button to edit data for a proteolytic enzyme.
Clicking the Property... button causes the Enzyme Property dialog box to appear.   Enter data in this dialog box to edit enzyme data.
You cannot change the name of an enzyme to the same name as that of any other registered enzyme.
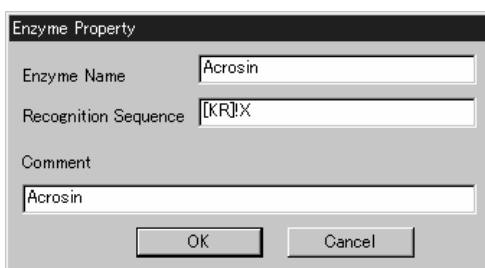
Enzyme Property dialog box

Item (parameter)	Description
Enzyme Name(NAME)	Enter the name of the proteolytic enzyme.
	The OK button is disabled if this field is blank.
Recognition Sequence(SITE)	Sequence recognized by the proteolytic enzyme.
	An amino acid sequence is represented in the single-character format with an exclamation mark (!) indicating a cut position.
	If there is more than one recognition sequence, a slash (/) is used as a delimiter.
	If there is more than one recognition amino acid (complex code), the acids are enclosed by brackets [ ].
	X indicates any amino acid.
	Example: [KR] ! X / AR ! X
	Identify KX, RX, and ARX and cut between K and X, R and X, and AR and X.
	The OK button is disabled in the following cases:
	For each sequence separated by a slash (/),
	1. There are more than one cut position (!).
	2. The data does not contain any amino acid characters.
	3. Any character other than A to Z, !, and [ ] is used.
	4. Brackets [ ] are nested in other brackets [ ].
	5. Brackets [ ] are not paired.
	6. Brackets [ ] contain X.
	7. Brackets [ ] contain no characters.
Comment	Displays comments for the proteolytic enzyme, if any.

Button	Description
OK button	Registers the proteolytic enzyme with the entered data.
	DNASIS MAX cannot register an enzyme if its modified name is already used for an existing enzyme. In such a case, you must change the name to register it.
Cancel button	Cancels the editing of new proteolytic enzyme data.

Errors that may occur when editing data

Duplicate enzyme name
For renaming an enzyme that is already registered, the dialog box appears.
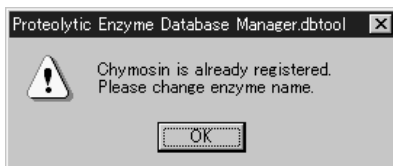
Click the OK button to return to the Enzyme Property dialog box.   Change the name of the enzyme and retry.

Importing Proteolytic Enzyme Data
In the main window, you can click the Import... button to import data for a proteolytic enzyme.
Clicking the Import... button causes the following dialog box to appear.   Select the file to import in this dialog box, and import data.
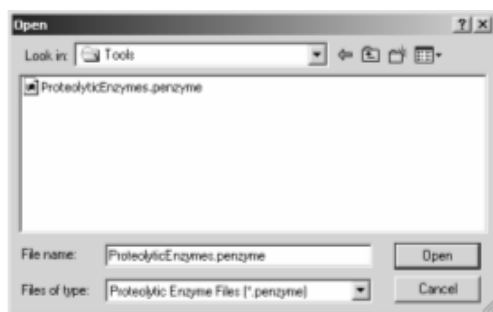
Errors that may occur when importing data

Non-proteolytic enzyme data
If any data in the file is not proteolytic enzyme data, the dialog box appears.
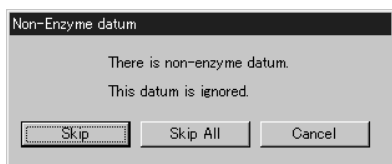
To skip that non-proteolytic enzyme data and continue processing, click the Skip button.
To skip all subsequent non-proteolytic enzyme data, click the Skip All button.
To cancel importing data, click the Cancel button.

Recognition site format error
If the data in the file specifies a recognition site in an invalid format, the dialog box appears.   To skip that invalid data and continue processing, click the Skip button.

To skip all subsequent invalid data, click the Skip All button.
To cancel importing data, click the Cancel button.
DNASIS MAX assumes data to be invalid in the following cases:
For each sequence separated by a slash (/),

1. There are more than one cut position (!).

2. The data does not have any amino acid characters.

3. Any character other than A to Z, !, and [ ] is used.

4. Brackets [ ] are nested in other brackets [ ].

5. Brackets [ ] are not paired.

6. Brackets [ ] contain X.

7. Brackets [ ] contain no characters.

No enzyme name
If the enzyme data in the file does not have a name, the dialog box appears.
To skip that enzyme data and continue processing, click the Skip button.
To skip all subsequent unnamed data, click the Skip All button.
To cancel importing data, click the Cancel button.
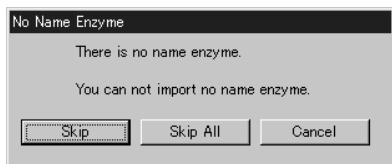
Duplicate enzyme name
If the enzyme data in the file has a name which is used for another existing enzyme, the dialog box appears.
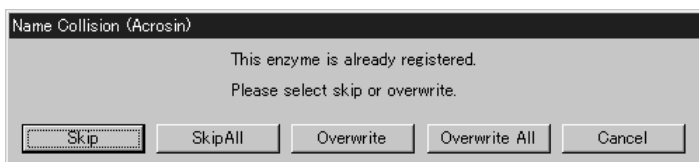
To skip that enzyme data and continue processing, click the Skip button.
To skip all subsequent duplicate named data, click the Skip All button.
To overwrite the existing enzyme data with the imported data, click the Overwrite button.
To overwrite all subsequent duplicate named enzyme data with the imported data, click the Overwrite All button.
To cancel importing data, click the Cancel button.

Exporting Proteolytic Enzyme Data

In the main window, you can click the Export... button to export data for the proteolytic enzyme.
Clicking the Export... button causes the following dialog box to appear.   Select the name of the file to which you want to export data in this dialog box, and export data.
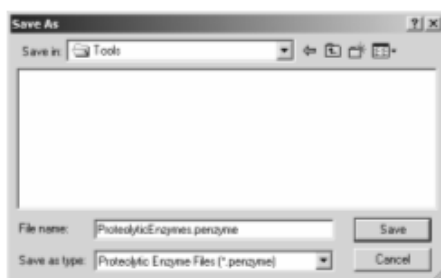
Errors that may occur when exporting data

Too long enzyme name
If the name of the proteolytic enzyme exceeds 255 characters, the following dialog box appears:

DNASIS MAX does not export this data and proceeds to exporting next data.

No enzyme name
If the selected proteolytic enzyme does not have a name, the dialog box appears.

DNASIS MAX does not export this data and proceeds to export the next data.

5.11 BLAST Search Dedicated Database

Use this window to create and manage sequence databases dedicated to BLAST search.

Window description
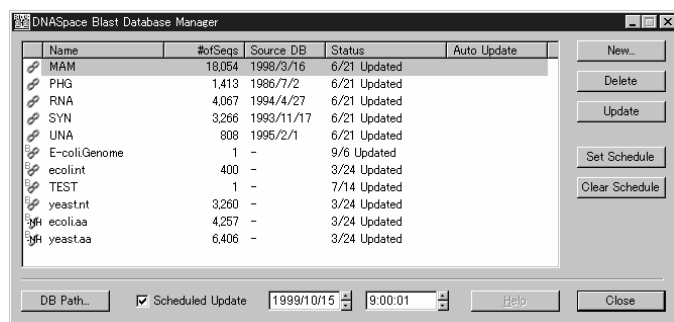
Icon	Description
	This icon indicates that the database has been converted from a sequence database storing public DNA sequences (mainly GenBank).
	This icon indicates that the database has been converted from a sequence database storing public amino acid sequences.
	This icon indicates that the database has been converted from a sequence database storing in-house DNA sequences (such as the experimental data available).
	This icon indicates that the database has been converted from a sequence database storing in-house amino acid sequences.
	This icon indicates that the database stores DNA sequences dedicated to BLAST search. You can directly copy a file, for example, created by format db of the NCBI tool kit.
	This icon indicates that the database stores amino acid sequences dedicated to BLAST search. You can directly copy a file, for example, created by format db of the NCBI tool kit.

Item	Description
Name	Displays the name of the database.
# of Seqs	Displays the number of sequence data items stored in the database.
Source DB	Source DB Displays the date on which the source sequence database was updated last.
Status	Displays "Empty" or the last updated date.
Auto Update	Displays "Scheduled" if automatic update is specified.

*Refer to "Select Sequence Database dialog box" in "5.11 Blast Search Dedicated Database".

Button	Description
New	Converts a database registered in the Sequence DB Manager to a database dedicated to BLAST search. The Select Sequence Data Base dialog box* appears. However, this button only creates an empty database without actually converting the database. You must subsequently update it.
Delete	Deletes the database.
Update	Immediately updates the database. Actually, DNASIS re-creates a database by converting all entries of the source sequence data.
Set schedule	With a database selected, clicking this button sets an update schedule.
Clear schedule	Clears the update schedule settings.
DB Path	Allows you to specify the path of the directory to store the dedicated database for BLAST searches.
Scheduled Update	Entering a date and time and checking this check box causes DNASIS to automatically update the database on the specified date and time. You can press the Delete key in the date field to clean the field. In that case, DNASIS will update the database every day at the specified time. However, you cannot use that function for a dedicated database for BLAST searches.
Help	Displays online help.
Close	Closes the BLAST DB Manager.

Select Sequence Database dialog box

In the Sequence DB Manager, select the data you want to convert and click Make.   DNASIS registers the selected data with the BLAST DB Manager.

In the BLAST DB Manager, clicking the New button causes the following dialog box to appear:
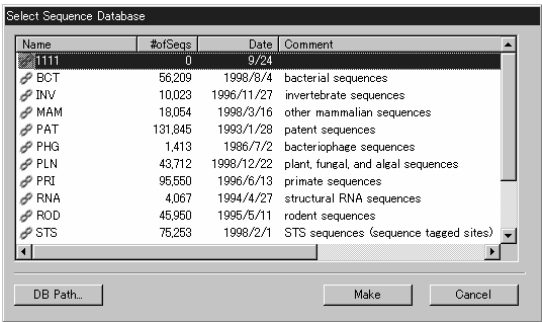
Item	Description
Name	Displays the name of the database.
# of Seqs	Displays the number of sequence data items stored in the database.
Date	Displays the date when the database was updated last.
Comment	Displays comments, if any.

Button	Description
DB Path	Allows you to specify the path of the directory to store the sequence database.
Make	Creates a dedicated database for BLAST searches from the database.
Cancel	Returns to the previous screen.

Chapter 6 Create Plasmid Maps

6.1 About Creating Plasmid Maps

It is possible to create a plasmid map of a selected sequence in Sequence View. A plasmid map is represented as a circle with the name and length of the sequence located in the center. Restriction enzymes are put on the circumference of the plasmid based on position. Annotations are displayed using arrows based on the start and end positions.
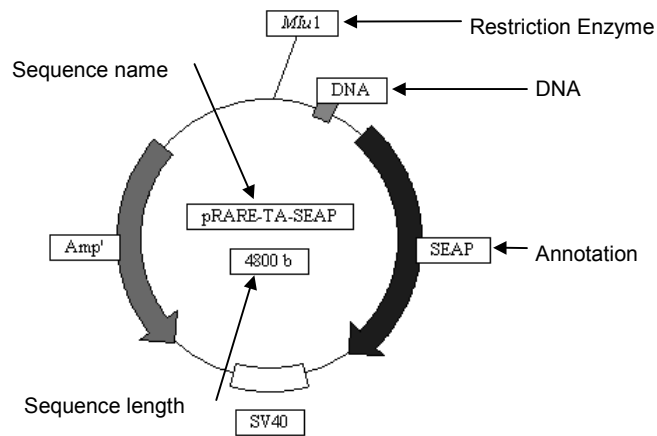
Plasmid maps can be edited by adding or changing plasmid figures such as restriction enzymes, annotations and DNA, or deploying normal figures such as rectangles and helices.

6.2 Create a Plasmid Map

1. Select a sequence in Sequence View.

2. Select Option from the Analysis Category in Analysis Button View, and click Plasmid View.
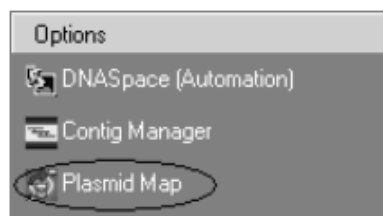
3. The plasmid map of the selected sequence will appear.
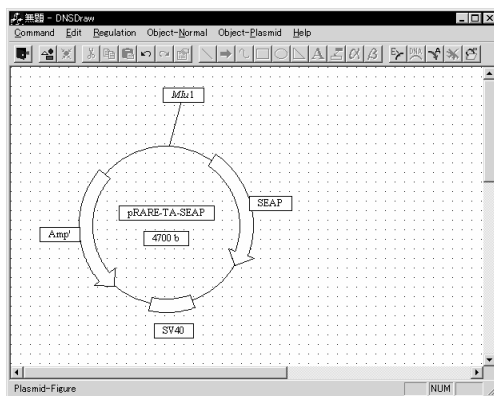
6.3 Map Editing Window
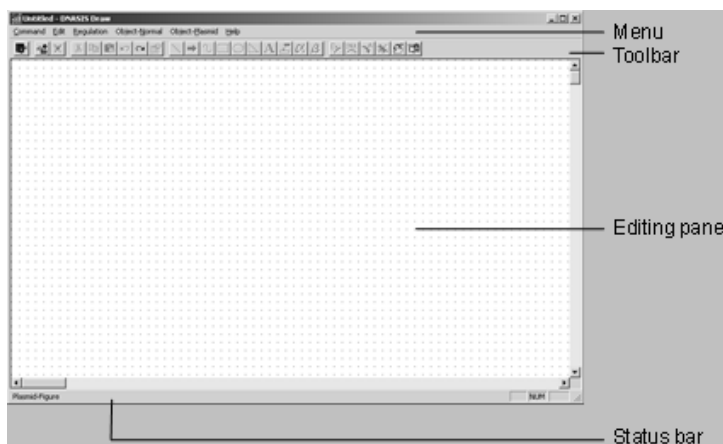
6.3.1 Menu

Command Menu	Description
Export Template	Exports the current status to a template.
Import Template	Reads in and displays figures from a template.
Preview	Displays a print preview.
Print	Starts printing.
Normal-Figure	For inputting and editing normal figures (Normal mode).
Plasmid-Figure	For inputting and editing plasmid figures (Plasmid mode).
Exit	Closes the Figure Editing window.

Edit menu	Description
Cut	Cuts the selected object.
Copy	Copies the selected object.
Paste	Pastes the cut or copied object.
Undo	Cancels the previous operation.
Redo	Restores the canceled operation.
Properties	Displays the properties of the selected object.

Regulation menu	Description
Spin left	Rotates the selected object 90% in a counterclockwise direction.
Spin right	Rotates the selected object 90% in a clockwise direction.
Spin free	Rotates the selected object any given angle.
Reverse Horizontal	Inverts the selected object.
Reverse Vertical	Reverses the selected object.
Bring to Front	Moves the selected object to the front.
Send to Back	Moves the selected object to the back.
Bring Forward	Moves the selected object forward.
Send Backward	Moves the selected object backward.
Group	Groups the selected objects.
Ungroup	Ungroups selected objects.

Object-Normal menu	Description
Line	Draws straight lines.
Arrow	Draws arrows.
Curve	Draws curved lines.
Rectangle	Draws rectangles.
Ellipse	Draws ellipses.
Polygon	Draws polygons.

Text	Creates text areas.
Label	Creates balloon texts.
Spiral Type alpha	Draws spirals of Spiral Type alpha helix.
Spiral Type beta	Draws spirals of Spiral Type beta helix.

Object-Plasmid menu	Description
Add restriction enzyme	Adds restriction enzymes.
Insert DNA by enzyme	Adds DNA to the positions of the selected restriction enzymes.
Annotation	Adds annotations to plasmid regions.
Delete Object	Deletes selected restriction enzymes, DNA or annotations.
Read file	Imports external files.
Alignment of Label	Selected: When editing, the restriction enzyme is automatically realigned.
	Not selected: When editing, the restriction enzyme position does not change.

Help menu	Description
Version	Displays version information.

6.3.2 Toolbar

Icon	Description
	Closes the Map Editing window. The same as Command > Exit in the menu.
	For inputting and editing normal figures (Normal Mode). The same as Command > Normal-Figure in the menu.
	For inputting and editing plasmid figures (Plasmid Mode). The same as Command > Plasmid-Figure in the menu.
	Cuts the selected object. The same as Edit > Cut in the menu.
	Copies the selected object. The same as Edit > Copy in the menu.
	Pastes the cut or copied object. The same function as Edit > Paste in the menu.
	Cancels the previous operation. The same as Edit > Undo in the menu.
	Restores the canceled operation. The same as Edit > Redo in the menu.
	Displays the properties of the selected object. The same as Edit > Properties in the menu.
	Draws straight lines. The same as Object-Normal > Line in the menu.
	Draws arrows. The same as Object-Normal > Arrow in the menu.
	Draws curved lines. The same as Object-Normal > Curve in the menu.
	Draws rectangles. The same as Object-Normal > Rectangle in the menu.
	Draws ellipses. The same as Object-Normal > Ellipse in the menu.
	Draws polygons. The same as Object-Normal > Polygon in the menu.
	Creates text areas. The same as Object-Normal > Text in the menu.
	Creates balloon texts. The same as Command > Label in the menu.
	Draws spirals of Spiral Type alpha helix. The same as Object-Normal > Spiral Type alpha in the menu.
	Draws spirals of Spiral Type beta helix. The same as Object-Normal > Spiral Type beta in the menu.
	Adds restriction enzymes. The same as Object-Plasmid > Add restriction enzyme in the menu.
	Adds DNA to the positions of the selected restriction enzymes. The same as Object-Plasmid > Insert DNA by enzyme in the menu.
	Adds annotations to plasmid regions. The same as Object-Plasmid > Annotation in the menu.
	Deletes selected restriction enzymes, DNA, or annotations. The same as Object-Plasmid > Delete Object in the menu.

Icon	Description
	Reads in external files. The same as Object-Plasmid > Read file in the menu.
	Auto align or not control. The same as Alignment of Label in the menu.

6.3.3 Status Bar

Displays the current edit mode (Normal or Plasmid).

6.4 Draw in Plasmid Mode

In Plasmid Mode, plasmid maps can be drawn and edited. To create or edit in Plasmid Mode, select Command >

Plasmid-Figure in the menu, or click ⌖ on the Toolbar. When drawing in Plasmid Mode, creating and editing normal figures is not allowed.

When there is no plasmid circle in the editing area, such as after deleting one, click ⌖ to open a dialog to create a new plasmid circle. The operation is as follows.

1. Click ⌖ when there is no plasmid circle in the editing area. The Plasmid Component dialog will appear.
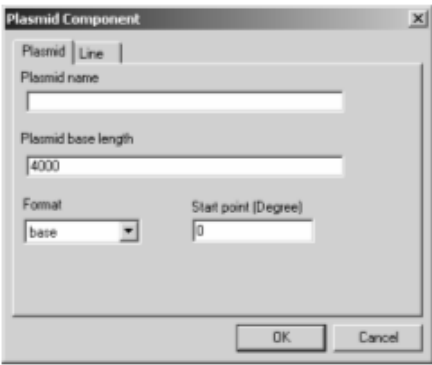
Item	Description (Initial setting)
Plasmid Name	Specifies the plasmid name using up to 50 single-byte characters to show in the center of the circle. (-)
Plasmid base length	Specifies the plasmid base number. (4000) Minimum value: 100 Maximum value: 99999
Format	Selects the displaying format of plasmid base number. (base)
Start point	Specifies the start position of the base sequence from 0 to 359. (0)

When tabs other than the Plasmid Tab are selected, it is possible to set the line type or thickness.

2. Input necessary information and click OK. A plasmid circle will appear. However, when "unknown" is specified in Format, the base sequence number will not display.
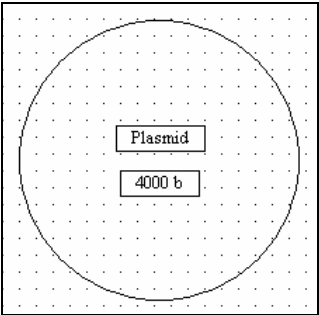
Add Restriction Enzyme

1. Click ⧽ on the Toolbar, and click a point on the plasmid circle circumference. The Restriction enzyme Component dialog will appear.
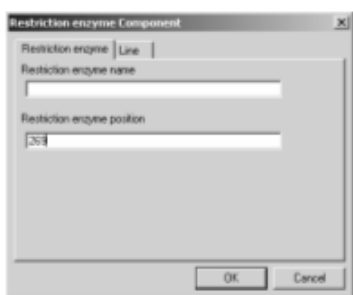
Item	Description (Initial setting)
Restriction enzyme name	Specifies the restriction enzyme name using up to 15 single-byte characters. (-)
Restriction enzyme position	Specifies the position of the restriction enzyme in a plasmid base. (Specified position)
	Minimum value: 1
	Maximum value: Plasmid base length number.

When tabs other than the Restriction enzyme Tab are selected, it is possible to change the line types and thickness of the drawn line.

2. Input necessary information and click OK. The restriction enzyme will be added, and the restriction enzyme name will appear (the first three characters are in italic).
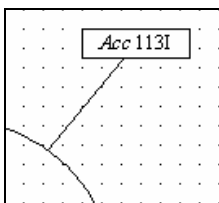
Inserting DNA

1. Select one or two restriction enzymes* of the part to insert, and click on the Toolbar. The DNA Component dialog will appear. However, DNA cannot be inserted into areas overlapping with existing DNA or annotations.

   *In order to select two enzymes, click the second enzyme while pressing the Shift key. When two enzymes are selected, DNA will be inserted into the one with shorter spacing.

Item	Description (Initial setting)
DNA name	Specifies the DNA name using up to 50 single-byte characters. (-)
Insert start position	Specifies the start position of the insertion region. (Specified position)
	Minimum value: 1
	Maximum value: Plasmid base length number.
Insert end position	Specifies the end position of the insertion region. (Specified position)
	Minimum value: 1
	Maximum value: Plasmid base length number.
DNA base length	Specifies the number of bases to inserting DNA. This item is required. (1)
Direction	Selects the insertion direction. (Clockwise/forward)
	Clockwise/forward
	Counterclockwise/backward
	Non-direction

When tabs other than the DNA Tab are selected, it is possible to change the line type and arrow thickness.

2. Input necessary information and click OK. The DNA will be added, and the DNA name will appear outside of the circle. The specified restriction enzymes and the restriction enzymes in the specified area will be deleted. Also, the difference between the number of DNA bases to insert and the number of bases of restriction enzymes in the specified region will be added to the total number of plasmid base number and the position of objects after the insertion point. When non-direction is specified, or the length is shorter than the arrowhead, the arrowhead does not appear.
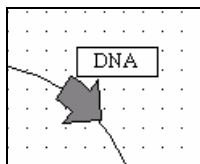
Adding an Annotation

1. Click on the Toolbar, and drag clockwise on the plasmid circumference from the start to the end position. The Annotation Component dialog will appear.
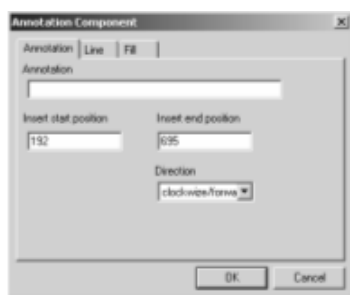
Item	Description (Initial setting)
Annotation	Specifies an annotation of up to 50 single-byte characters. (-)
Insert start position	Specifies the start position of the inserting area. (Specified position)
	Minimum value: 1
	Maximum value: Plasmid base length number.
Insert end position	Specifies the end position of the inserting area. (Specified position)
	Minimum value: 1
	Maximum value: Plasmid base length number.
Direction	Selects the adding direction. (Clockwise/forward)
	Clockwise/forward
	Counterclockwise/backward
	Non-direction

When tabs other than the Annotation Tab are selected, it is possible to change the line type and arrow color.

2. Input necessary information and click OK. The annotation will be added. Also, when the arrow overlaps with the existing arc arrow, an arc arrow will be created outside the circumference.
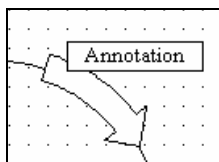
Change the Plasmid Circle

The size of a plasmid circle can be changed and moved by the mouse. Drag the handle to change the size, and the positions and sizes of the relevant figures also change. Also, when moving a circle, the relevant figures are moved together. The plasmid name and the text area for the base sequence can also be changed or moved, but the text area cannot be deleted.

Figures can also be changed by changing the plasmid properties. The operation is described below:

1. Select a plasmid circle, and click on the Toolbar. The Plasmid Component dialog will open.
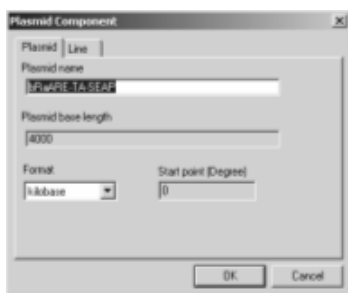
Item	Description
Plasmid Name	Specify the name to show in the center using up to 50 single-byte characters.
Plasmid base length	Displays the number of plasmid bases. This item cannot be changed.
Format	Selects the display format of plasmid bases.
Start point	Displays the start position of the base sequence. This item cannot be changed.

When tabs other than the Plasmid Tab are selected, it is possible to change type and thickness of the line.

2. Input necessary information and click OK. The plasmid circle will be changed.

Change Restriction Enzyme

It is possible to move or change the size of the restriction enzyme name text area by using the mouse. Even when this text area is moved, the positions (lines that show the links on the circumference) of the restriction enzymes won't change.

Figures can also be changed by changing the restriction enzyme properties. The operation is described below.

1. Select a restriction enzyme, and click 📄 on the Toolbar. The Plasmid Component dialog will open.
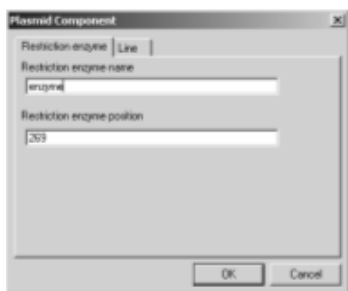
Item	Description
Restriction enzyme Name	Specify the name for the restriction enzyme up to 15 single-byte characters.
Restriction enzyme position	Specify the location of the restriction enzyme in the plasmid bases.
	Minimum value: 0
	Maximum value: Plasmid base length number.

When tabs other than the Restriction enzyme Tab are clicked, it is possible to change the type and thickness of the drawn line.

2. Input necessary information and click OK. The restriction enzyme will be changed. Also, when the Restriction enzyme position is changed, the figures will be moved to the corresponding positions.

Change the DNA

It is possible to move or change the size of the DNA name text area by using the mouse.

Figures can also be changed by changing the DNA properties. The operation is described below:

1. Select a DNA, and click 📄 on the Toolbar. The Plasmid Component dialog will open.
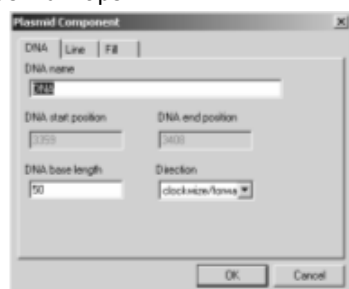
Item	Description
DNA Name	Specify the DNA name using up to 50 single-byte characters.
DNA start position	Displays the start position of the DNA. This item cannot be changed.
DNA end position	Displays the end position of the DNA. This item cannot be changed.
DNA base length	Specifies the number of bases of the DNA.
	Minimum value: 1
	Maximum value: Plasmid base length number.
Direction	Selects the direction of the DNA.
	Clockwise/forward
	Counterclockwise/backward
	Non-direction

When tabs other than the DNA Tab are clicked, it is possible to change the arrow line type and color.

2. Input necessary information and click OK. The DNA will be changed. Also, when the DNA base length is changed, the end position will automatically be changed. Additionally, when non-direction is specified, or the length is shorter than the arrowhead, the arrowhead does not appear.

Change Annotation Length

Drag the handle at the annotation start or end position to change the length.

Change the Length
Drag the handle at the annotation start or end position to change its length.
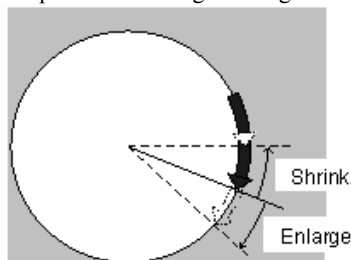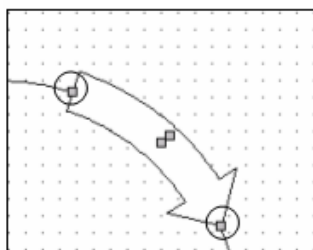
Move the Position along the Circumference
Drag the handle in the center of the annotation to move along the circumference.
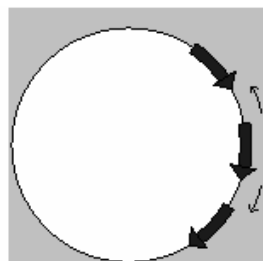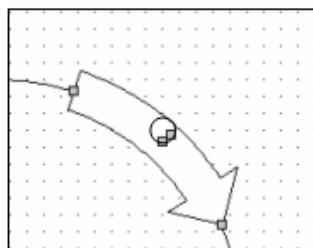
Move the Position Perpendicular to Circumference
Drag the handle next to the center of the annotation to move in a direction perpendicular to the circumference. The size changes automatically, when moving this way.
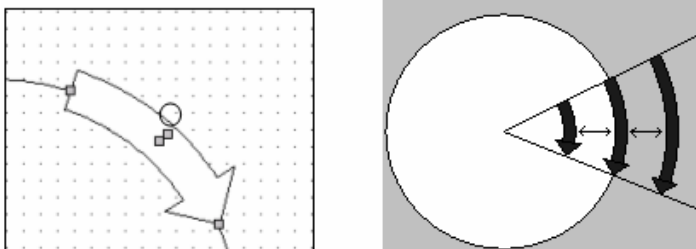
Figures can also be changed by changing the annotation properties. The operation is described below.

1. Select an annotation and click [icon] on the Toolbar. Or change the annotation length and move along the circumference by using the mouse. The Plasmid Component dialog will appear.

Item	Description
Annotation	Specifies the annotation using up to 50 single-byte characters.
Annotation start position	Specifies the start position of the annotation. Minimum value: 1 Maximum value: Plasmid base length number.
Annotation end position	Specifies the end position of the annotation. Minimum value: 1 Maximum value: Plasmid base length number.
Direction	Selects the direction of the annotation. Clockwise/forward Counterclockwise/backward Non-direction

When tabs other than the Annotation Tab are clicked, it is possible to change the arrow line type and color.

2. Input necessary information and click OK. The annotation will be changed. When non-direction is specified, or the length is shorter than the arrowhead, the arrowhead does not appear. When the same value is designated for the Annotation start position and the Annotation end position, the annotation is represented as a line.

Delete Objects

It is possible to delete selected objects (plasmids, restriction enzymes, DNA, and annotations). When a plasmid is deleted, the relevant restriction enzymes, DNA, and annotations are all deleted. When a DNA is deleted, the base number of the DNA will be subtracted from the total base number of the plasmid and the object position after the delete position.

To delete an object, select it and click [icon] on the Toolbar.

Additionally, deleting only the text area of an object is not allowed.

Import a File

You can import external files in the dmp format by selecting Export in the File menu of the main window.

1. Click [icon] on the Toolbar. The dialog below appears:

2. Select a file and click Open. If there is a plasmid map being created, the following message, saying that the plasmid map being created will be destroyed, will appear.
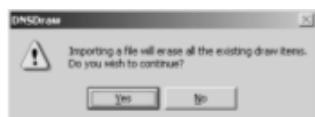
3. Click Yes to display the dialog below.
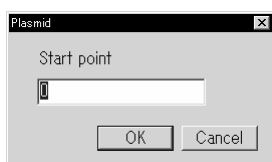
Item	Description (Initial setting)
Start point	Specifies the start position of the base sequence in the range from 0 to 359. (0)
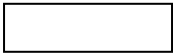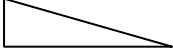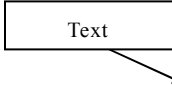
4. Specify the start position and click OK. A plasmid map will appear in the editing area. The plasmid map already on display will be overwritten by the one made from the imported file, but the figures edited in Normal Mode will display unchanged.

6.5 Drawing in Normal Mode

In Normal Mode normal figures such as lines, arrows, rectangles, and spiral diagrams can be drawn and edited. To

create or edit figures in Normal Mode, select Command > Normal-Figure, or click on the Toolbar.
While drawing in Normal Mode, plasmid figures cannot be created or edited.

Add Normal Figures

To draw normal figures, click the icon for normal figures, and drag from the starting point to the endpoint of a figure.
The following normal figures can be drawn.

Type	Object
Line	
Arrow	
Curve	
Rectangle	
Ellipse	
Polygon	
Text	Text
Label	Text

Add Spirals

Two types of spiral can be drawn.

Spiral Type alpha

1. Click on the Toolbar, and drag the line that will be the center of the spiral while editing from the start position to the end position.

2. A spiral will be drawn.

Spiral Type beta
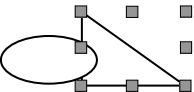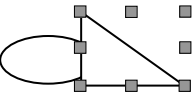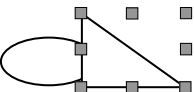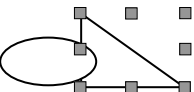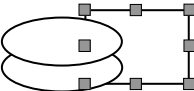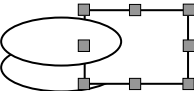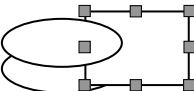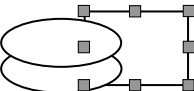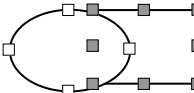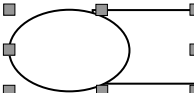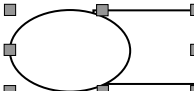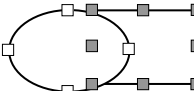
The operation is described below.

1. Click on the Toolbar, and drag the line that will be the center of the spiral while editing from the start position to the end position.
2. A spiral will be drawn.

Adjust a Figure

It is possible to adjust (such as rotate, reverse) normal figures and spirals. The following adjustments are possible.

Type	Before Adjustment	After Adjustment
Spin left		
Spin right		
Spin free		
Reverse Horizontal		
Reverse Vertical		
Bring to Front		
Send to Back		
Bring Forward		
Send Backward		
Group		
Ungroup		

Change a Figure

The size or position of normal figures and spiral can be changed by using the mouse.
Figures can also be changed by changing the figure properties. The operations to change normal figures and spirals are described below:

Change a Normal Figure

By changing the property of a normal figure, the thickness and line color can be changed. To change the properties of a normal figure, select it, and click 🖼 on the Toolbar. When the dialog appears, enter the items to change.

Change a Spiral

By changing the property of a spiral, features such as color can be changed. The operation is described below.

1. Select a spiral and click 🖼 on the Toolbar. The Plasmid Component dialog will appear. Click the Spiral Tab.
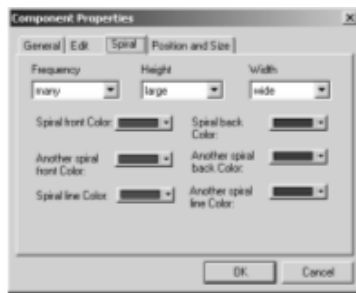
Item	Description
Spiral front Color	Selects the surface color of the spiral.
Spiral back Color	Selects the backside color of the spiral.
Another Spiral front Color	Selects the surface color of the other spiral.
Another Spiral back Color	Selects the backside color of the other spiral.
Spiral line Color	Selects the line color of the spiral.
Another Spiral line Color	Selects the line color of the spiral.
Frequency	Displays the frequency value of the spiral.
Height	Displays the frequency height of the spiral.
Width	Displays the band width of the spiral.

When tabs other than the Spiral Tab are clicked, it is possible to change the other properties.

2. Input the necessary information and click OK. The spiral will be changed.

6.6 Printing Figures

Plasmid maps can be printed.

1. Select Command > Print in the menu.

2. The plasmid map will be printed.

6.7 Working with Templates

Data in the process of creation can be exported to templates, or stored templates can be imported and displayed.

Export a Template

1. Select Command > Export Template in the menu. The dialog will appear.
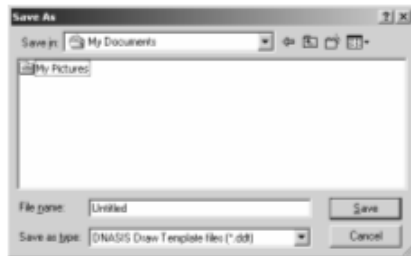
2. Type the file name, and click Save to export the file to a template.

Import a Template

1. Select Command > Import Template in the menu. The dialog will appear.

2. Specify the file name, and click Open and the template will be imported.

6.8 Exit Plasmid Map Drawing

Click the on the Toolbar. The editing window will close.

Chapter 7 Tutorial

7.1 Before Starting the Tutorial

7.1.1 About Installation

Using the Tutorial requires you to have sample data, which you can install from the Sample Database using the following procedures.

From the "Choose Destination Location" window, specify a location where you want to install the Sample Database (which is installed in C: \HSK_DB for the initial setting).
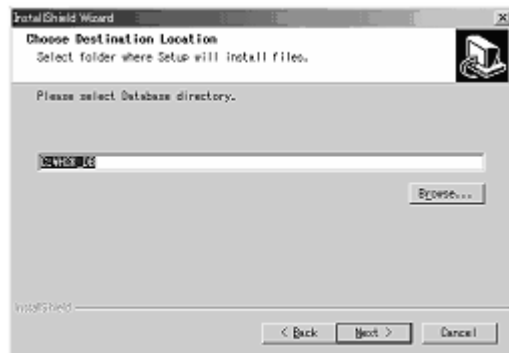
"Choose Destination Location" Window

The "Setup Type" window then appears.   Choose the Typical parameter as how to install it.
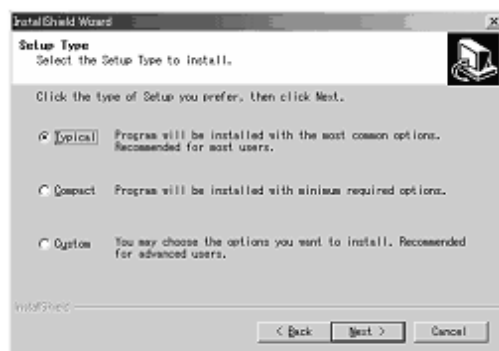
"Setup Type" Window

This Tutorial uses the file in the TutorialData folder in the Sample Database Installation Destination folder.   (For initial setting, it refers to the file contained in C: \HSK_DB\TutorialData.)

7.1.2 Data Used in the Tutorial

This section handles three versions of Tutorial, which are stored in the files listed below.

7.2 Open Reading Frame (ORF) Search	Tutorial1.fsa
7.3 Blast Search	Tutorial2.fsa
7.4 Vector Trimming	Tutorial3_1.abi, Tutorial3_2.fsa

7.1.3 Initial Settings

Some parts of this Tutorial require a connection to the Internet.   Depending on your network environment, you may need to configure the internet settings for DNASIS Max.

From the View menu in the Sequence Editor, select the Internet Options... item to display the settings window; alternatively, you can click the Internet Options button on the toolbar.

Our example here attempts to set the proxy server in the HTTP protocol for Web browsing.
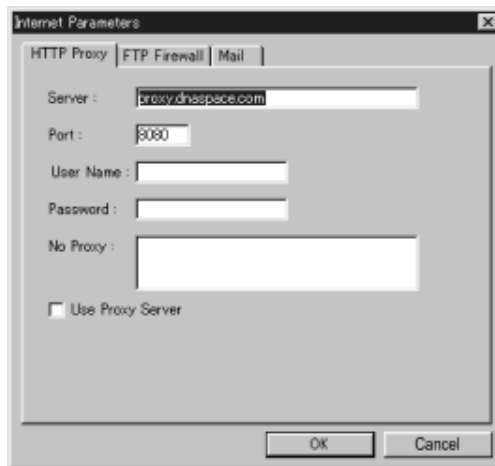
If you are not familiar with proxies, ask your network manager or refer to Internet Explorer for the settings.   In Internet Explorer, select Tools > Internet Options... from the menu.   Click the Connections tab, then click Settings... (for Dial-up connection) or LAN Settings... (for LAN connection).   If there is no check in the "Use a proxy server for this connection." (or "Use a proxy server for your LAN.") check box, you do not need to change the setting.   If there is a check in the box, check the "Use Proxy Server" item in DNASIS MAX and fill in the Server and Port information with the corresponding information in Internet Explorer.   If there are advanced settings, click Advanced... in Proxy Server of Internet Explorer and use that information.   Only when the Proxy Server requires user authentication, fill in the User Name and Password items.   You should leave these items blank if there is no need for user authentication.

7.2 ORF Search

This section deals with searching for Open Reading Frames (ORFs).   It also demonstrates searching for motifs in amino acid sequences that are generated by translating ORF data.   This section examines the following operations:
  • Starting DNASIS MAX
  • Entering the sequence
  • Searching for open reading frames
  • Translation
  • Searching for amino acid motifs

7.2.1 Starting DNASIS MAX

After clicking the Start button in Windows, select the following: Program, DNASIS MAX and then DNASIS MAX.

Alternatively, you can double-click the DNASIS MAX(.exe) icon ( ) in the DNASIS MAX installation destination folder.

When the program starts up a prompt dialog box will appear.
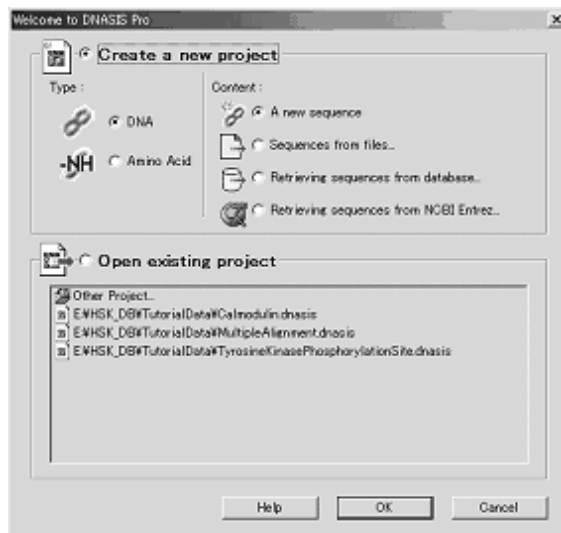
7.2.2 Using the Editor to Open Sequence Files

Select Create a new project from the prompt dialog. For Type select DNA and for Content select Sequences from files… then click the OK button. Specify Tutorial1.fsa from the dialog box that appears.
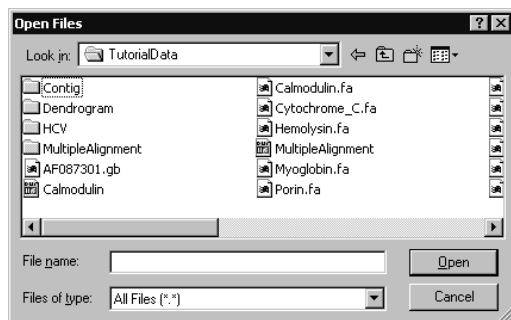
7.2.3 Running ORF Search

Select the "DNA - Search" group on the analysis button bar.
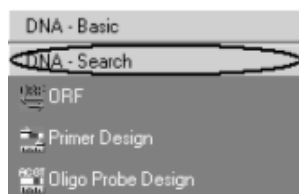
Click the ORF button ( ) and an Analysis dialog box will appear.
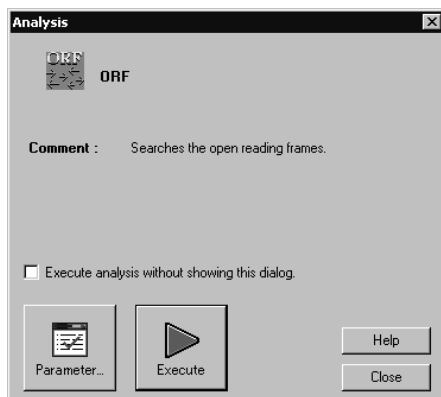
Click the Execute button to start an ORF search with 3 frames. When the analysis is finished the results will appear in map view and below the sequence in the sequence editor view.
Each ORF is indicated by an arrow.

Displaying the result of ORF search.
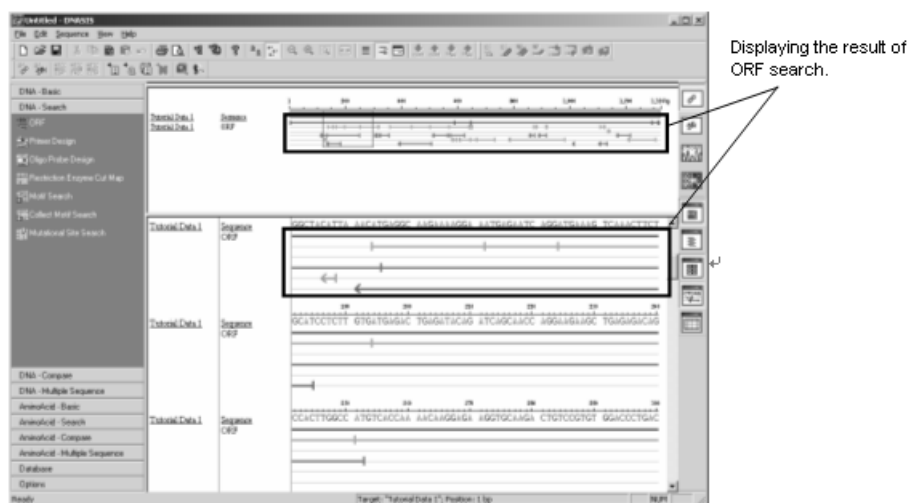
If You Want to Change the Codon Table

Click the ORF button ( ) and an Analysis dialog box will appear. Then click the Parameter button and a Parameter dialog box will appear.   Under the initial setting, "Universal" is found in the Codon Table of the Parameters item.   To change the Codon Table, select another table from the drop-down list and click the OK button.
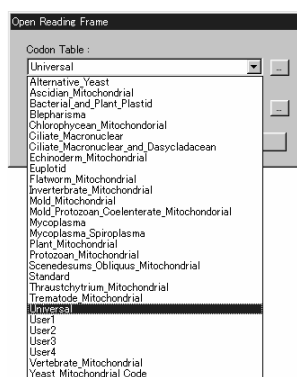
Similarly, you can specify the start codon with the Initial Codon parameter.
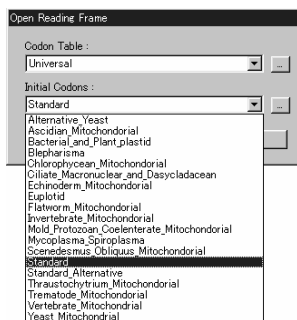
7.2.4 Running Translation

Select DNA – Bases from the analysis button bar then click the Translate button ( ) and an Analysis dialog box will appear. Click the Execute button to begin translation in the 3 forward reading frames.

Displaying the Result of Translation.
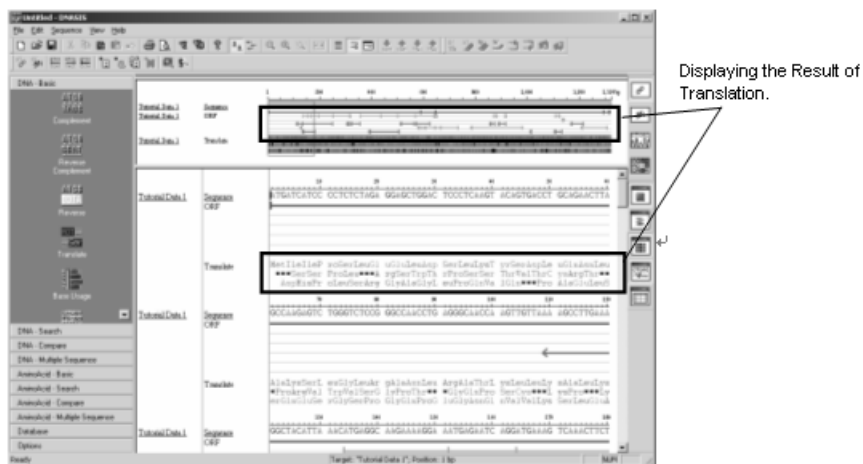
Changing the Translated Amino Acid Notation from Three Letter to One Letter

In the area for displaying the result of translation in the Sequence Editor view, right-click the mouse and select the Property... menu.

The Parameter Setting dialog box appears.   From the Amino Acid Symbols field, select One Letter and click the OK button.   As a result, the amino acid sequence changes to one-letter notation.
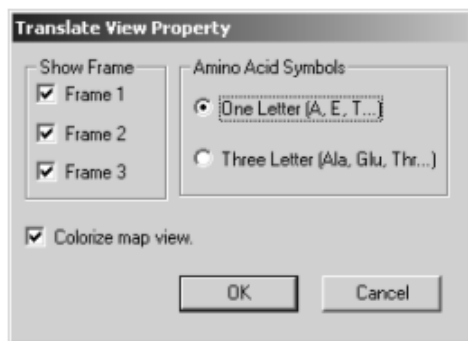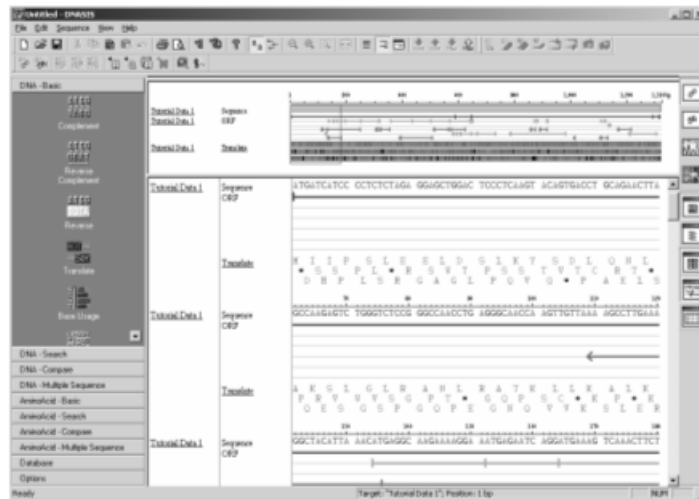
7.2.5 Displaying Only the Longest ORF

If there are too many ORFs to be read with ease, you can reduce the number of ORFs that are displayed at the same time. Right-click your mouse on the displayed ORF results and select the Show Settings menu.
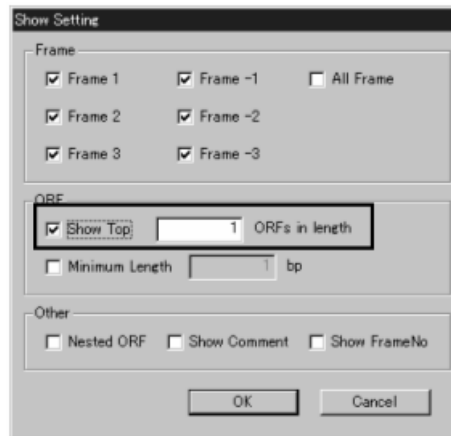
Place a checkmark for "Show Top [Number] ORFs in length" in the ORF field and fill in it with an appropriate number. When we select a number "n" here, only the "n" longest ORFs will be displayed.   In our example, we'll select "1" because we want to display only the longest ORF.

If You Want to Display the ORF List

In sequence view, select the sequence name and analysis name then click the Result List Dialog ( ) button.
If you want to store the ORF list, click the Save All button.   This allows the information displayed in.
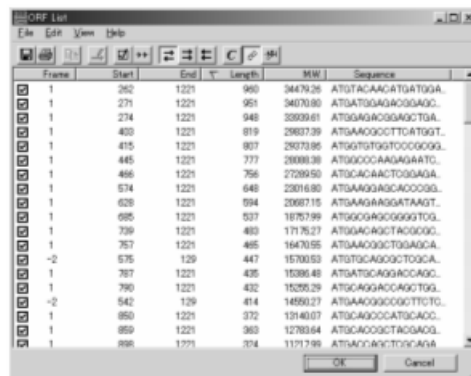
To save the ORF list, select Export from File in the menu.
The list will be stored in a text file containing the information separated by tabs.

7.2.6 Entering the Amino Acid Sequence for Selected ORFs into the Editor

Clicking the longest ORF in the Sequence View causes the amino acid sequence of the corresponding frame to be highlighted.   In our example, the amino acid sequence in the first reading frame is highlighted.
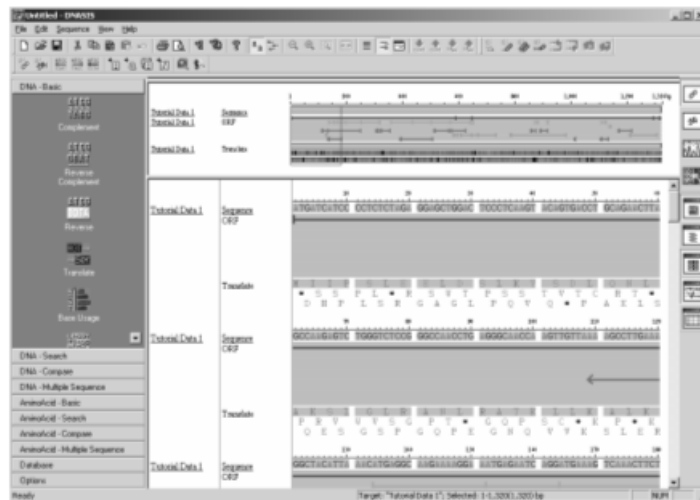
Click the Amino Acid Transfer ( ) button.   This moves you to the amino acid editing mode, in which the translated amino acid sequence is entered into the Editor.
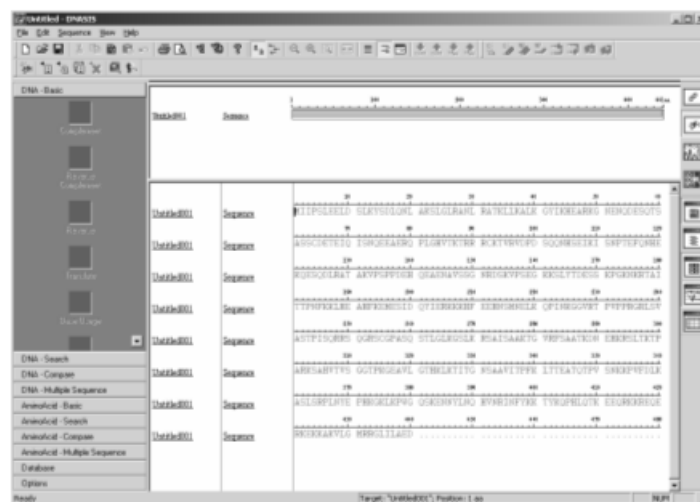
If You Want to View the Result of ORF Search
When a new amino acid sequence is created, the edit mode switches from " DNA sequence" to "amino acid sequence." Viewing the analysis result for DNA sequences requires you to go back to the nucleic acid sequence mode.   To do this, click the DNA Mode button ( ).   Click the Amino Acid Mode button ( ) if you want to edit the amino acid sequence once again.

7.2.7 Running Amino Acid Motif Search

After selecting the "Amino Acid - Search" group on the analysis button bar, click the Motif Search button ( ).
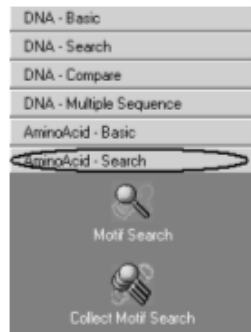
The search is performed according to the motifs that are registered with the amino acid motif database.   The results are displayed below the sequence in the Sequence View window and in the Map View window.
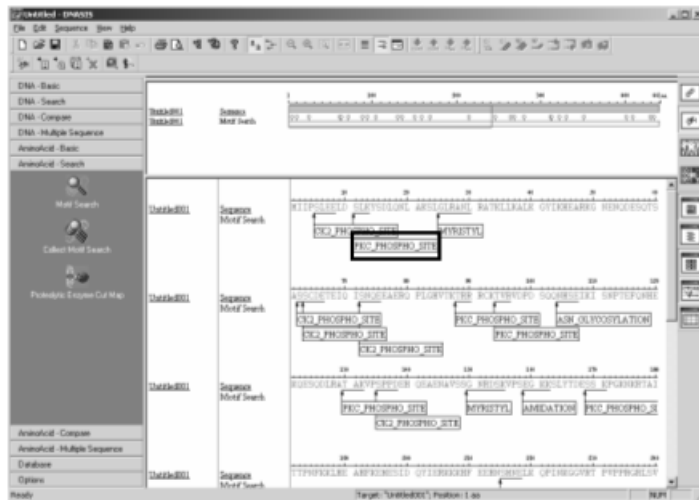
In the area displaying the result of amino acid motif analysis, double-clicking the motif name causes motif information to be displayed.
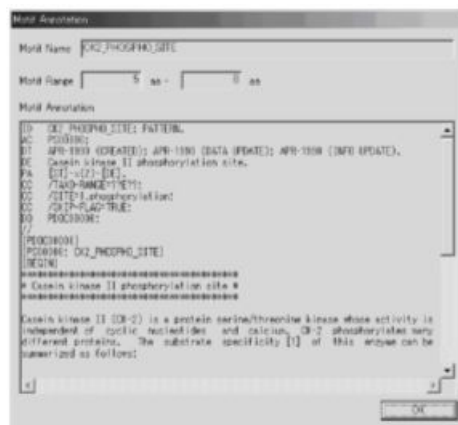
7.3 BLAST Search

This section demonstrates how to do a basic BLAST search using a local database.   It also demonstrates how to perform a multiple alignment using the result of a BLAST search.   The following operations are examined in this section:
- BLAST search
- Obtaining a GenBank file from NCBI
- Multiple alignment
- Adding annotations to the sequence

7.3.1 Starting DNASIS MAX

Refer to "Starting DNASIS MAX " in "7.2 ORF Search".

7.3.2 Using the Editor to Open Sequence Files

The Tutorial uses Tutorial2.fsa as its input sequence.   Select "Create a new project" from the initial dialog box. For Type select DNA and for Content select "Sequences from files…" then click the OK button. Specify Tutorial2.fsa from the dialog box that appears.   (For the location of tutorial data, refer to "7.1 Before Starting the Tutorial".)
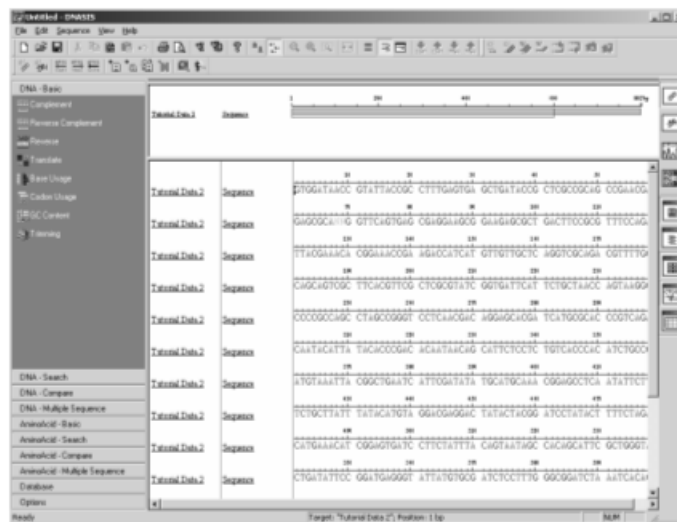
7.3.3 Specifying the Database as the Target of BLAST Search

It is necessary to set the search conditions before performing a BLAST search.   Select the "DNA - Compare" group on the analysis button bar.
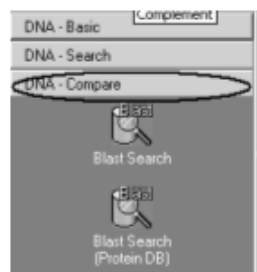
Click the BLAST Search button and an Analysis dialog box will appear. Then click the Parameter button, and the parameter setting dialog box for BLAST search will appear, showing the list of the various DNA sequence databases currently installed on the PC in the Nucleotide Database.   This tutorial will use the mammalian database (MAM), so make sure only MAM is selected and click the OK button.
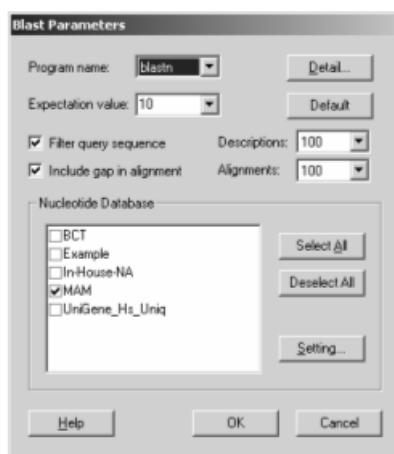
7.3.4 Running BLAST Search

Click the BLAST Search button and Analysis dialog box will appear. Clicking the Execute button starts BLAST search on the MAM database.   At the end of analysis, the result window appears.
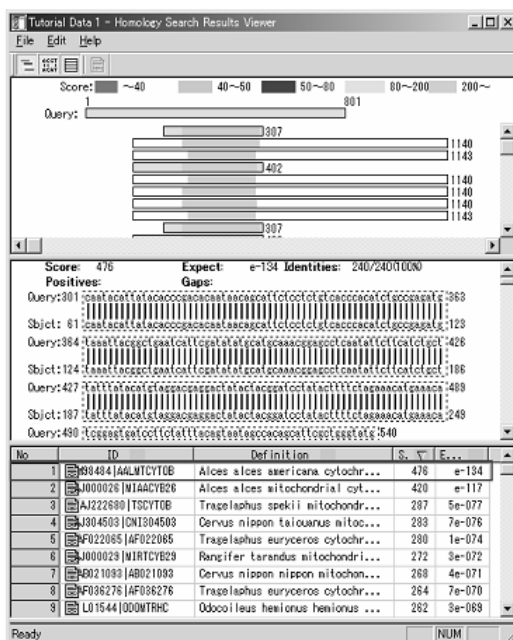
7.3.5 Using the Editor to Enter the Highest-Homology Sequence as a New Sequence from the Search Result Window

From the result list, select the hit with the greatest similarity (ID: M98484|AALMTCYTOB) and click the Get GenBank report button ( ) on the toolbar.
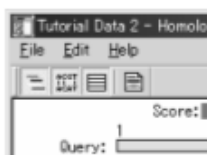

DNASIS Max will then attempt to acquire a file with the M98484 accession number in the GenBank format via the Internet using the Entrez system at NCBI.   If the attempt is successful, this sequence is added as a new sequence to the Editor.
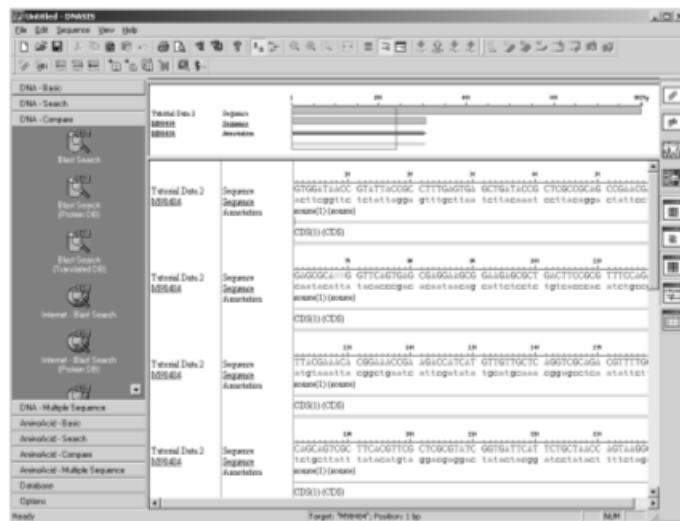
7.3.6 Running Multiple Alignment

After selecting the "DNA - Multiple Sequence" group on the analysis button bar, click the Multiple Alignment button and an Analysis dialog box will appear. Click the Execute button. An alignment between two sequences using the ClustalW method is performed, with the result being displayed in the Editor.
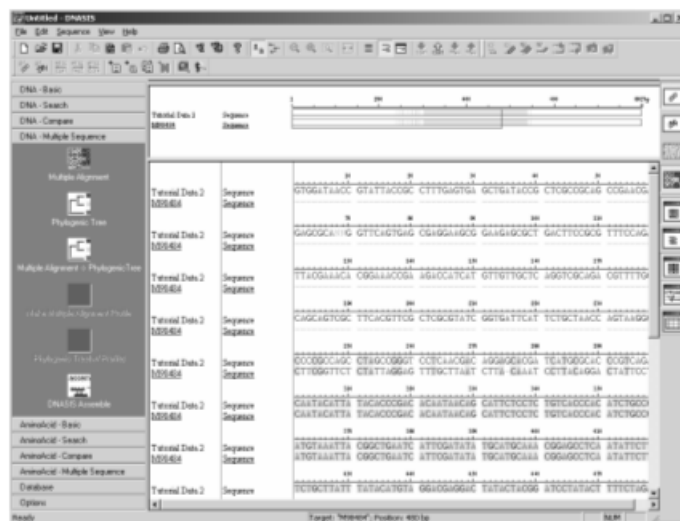
7.3.7 Adding Annotations to Similarities

A yellow background is shown on the portion having a match between two sequences.   Below we'll demonstrate how to add an annotation to the area of high homology in the 301bp to 540bp range in this example.
First, select the range from 301bp to 540bp using the mouse.
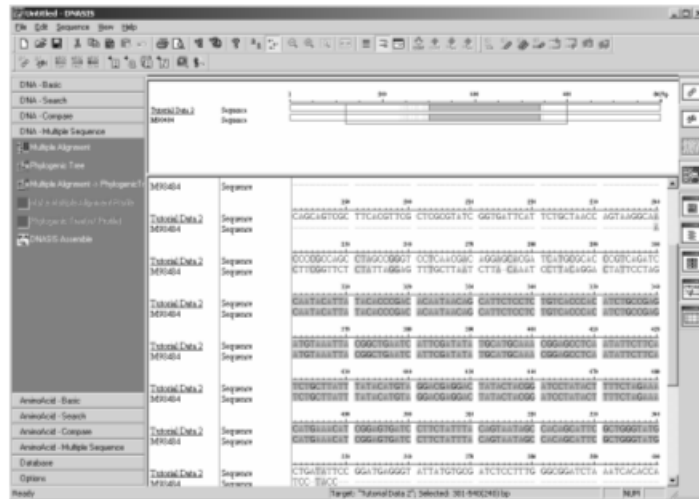
Since we are now in the alignment display mode, we need to return to the editing mode, that allows adding annotations. To this end, click the Sequence Mode ( ) button.

Click the New Annotation View button ( ) and then the New Annotation button ( ).

This displays the dialog box where you can add annotations.

Enter "Homology with M98484AALMTCYTOB" as the annotation name and click the OK button.   As a result, an annotation with the "Homology with M98484AALMTCYTOB" name is added.
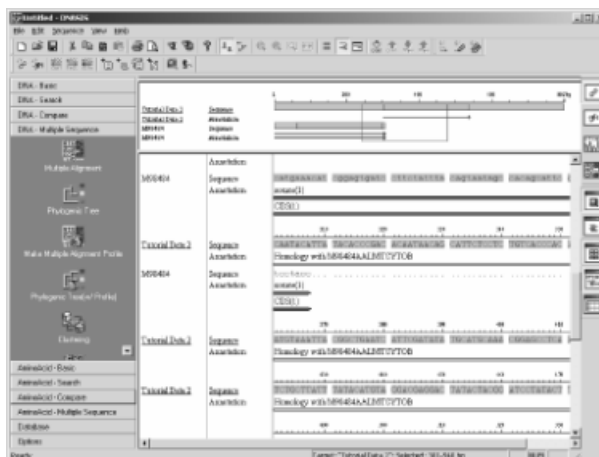
7.4 Vector Trimming

Vector trimming is used to find the vector sequence part from the output data of an automated DNA sequencer.   Vector sequence can then be masked.   You can also load a reference sequence and align it with the sequence based on the data from your sequencer, and display the two aligned sequences and the waveform data from your sequencer simultaneously.

- Waveform display
- Vector trimming
- Sequence masking
- Alignment with the reference sequence

7.4.1 Starting DNASIS MAX

Refer to "Starting DNASIS MAX " in "7.2 ORF Search".

7.4.2 Using the Editor to Open Sequence Files

The Tutorial uses Tutoria3_1.abi as its input sequence.   From the Sequence Editor's File menu, select Open and specify Tutoria3_1.abi.   (For the location of the tutorial data, refer to "7.1 Before Starting the Tutorial".)

7.4.3 Registering Vector Sequences with the Vector Database

Select the Database on the analysis button bar.

Clicking the Database button ( ) displays the vector database manager window.
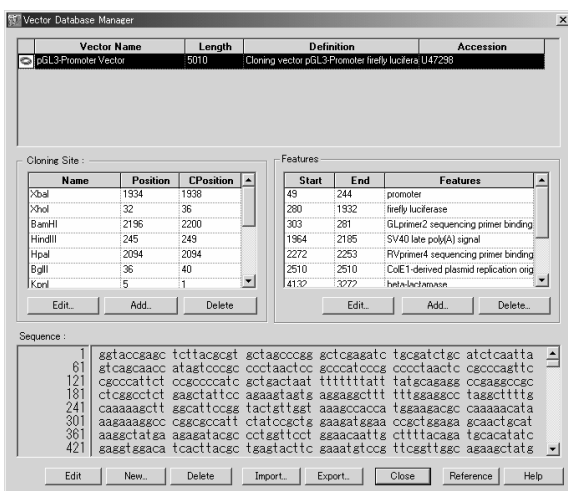
To register vector sequences for vector trimming, click the Import... button and specify the pSU2718.prm vector sequence file, which is located under the VectorData folder of the database installation destination.   (For a standard installation, you need to specify C: \HSK_DB\VectorData\pSU2718.prm.)
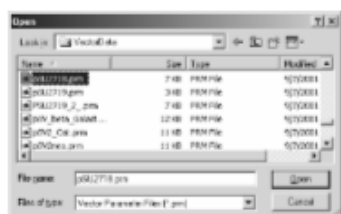
This imports the vector information so that a vector having the pSU2718 name is registered with the vector database.
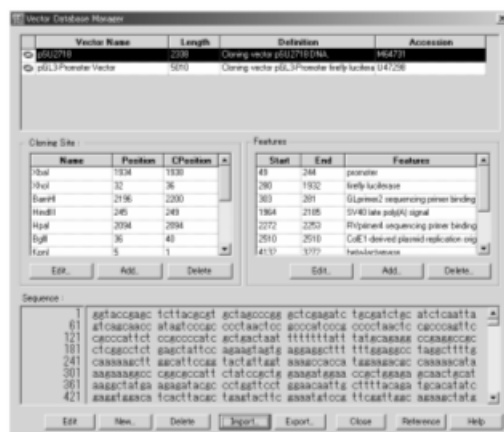

7.4.4 Carrying Out Vector Trimming

Select the "DNA - Basic" group on the analysis button bar.
Select the cloning vector that was used for sequencing and its cloning site.   Click the Vector and Low Quality Trim End

( ) button and an Analysis dialog box will appear.   Then click the Parameter button.
Select pSU2718 from the Vector Name list in the Trim Vector; and SmaI from the Cloning Site list.
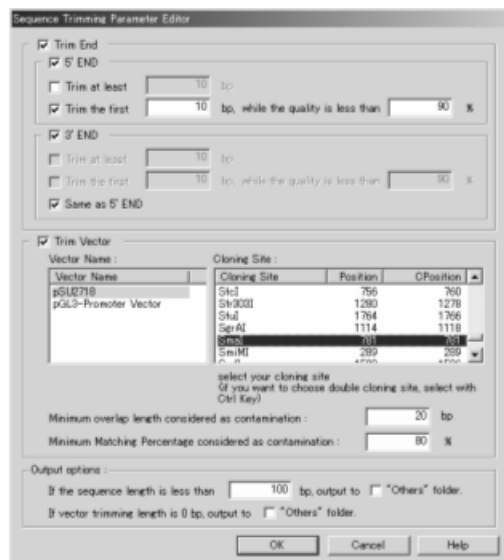
Click the OK button.
Click the vector and low-quality end trimming button.
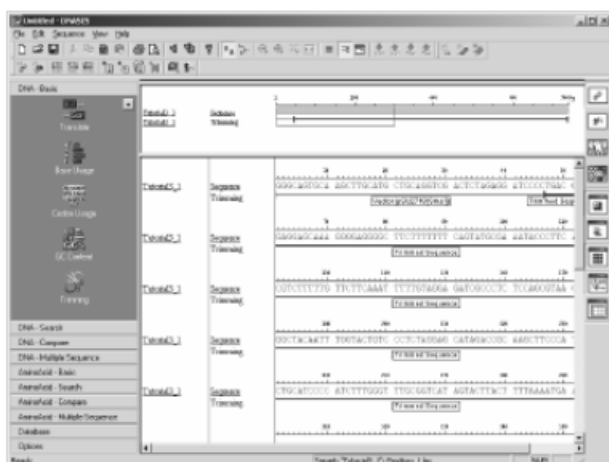
Once a vector sequence is found in the input sequence, Vector (pSU2718[SmaI]) as the vector sequence part is displayed below the sequence and Trimmed Sequence is displayed in the Insert filed.

7.4.5 Masking Vector Sequences

Select Vector (pSU2718[SmaI]) in the result of the vector and low-quality end trimming button to highlight the vector sequence.
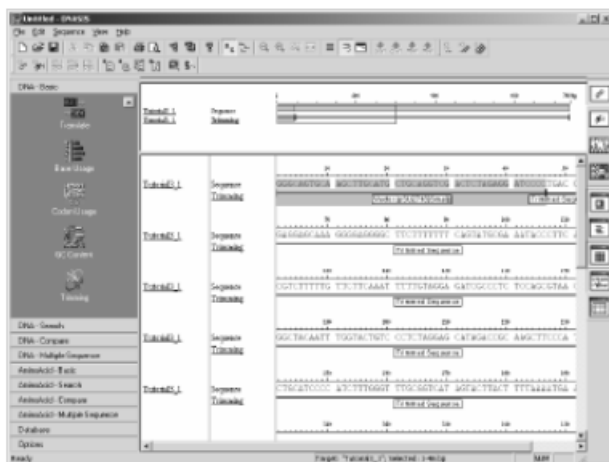
If, under this condition, you click the Mask button ( ) on the toolbar, the vector sequence is converted to N.
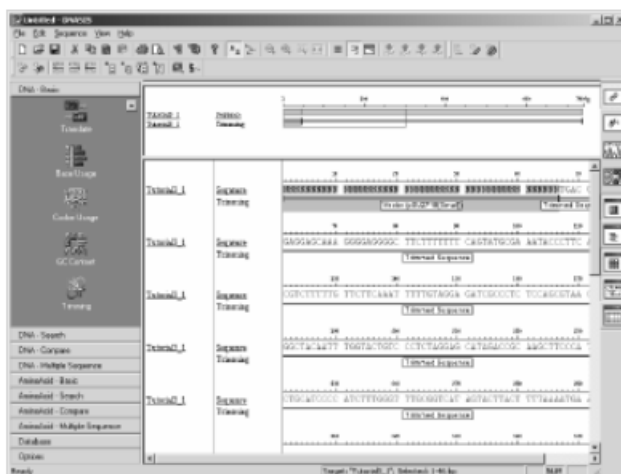
7.4.6 Switching to Waveform Display

Click the Trace View mode button ( ) to switch over to the waveform display mode.

7.4.7 Specifying the Reference Sequence

To find the mutation part, obtain "wild-type" reference sequences.

Click the Import Alignment Sequence button (⬓) and specify Tutorial3_2.fsa to obtain the reference sequence.   (For the location of tutorial data, refer to "7.1 Before Starting the Tutorial".)

7.4.8 Alignment with the Reference Sequence

Click the Show Alignments button ($\substack{A--T \\ ACGT}$) to align the input and reference sequences.

DNASIS Max uses the Clustal W method to align the two sequences.   The reference sequence is then displayed aligned with the input sequence and the corresponding waveform data for the input sequence.

Because any location having different bases between two sequences is highlighted, you can at a glance identify where a mutation has taken place.   In our example of the Tutorial, we can see mutations at the 59bp and 74bp locations.

Index